B1

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2002 (10.05.2002)

PCT

(10) International Publication Number
WO 02/36613 A2

(51) International Patent Classification[7]: C07K 14/00

(21) International Application Number: PCT/EP01/12778

(22) International Filing Date:
5 November 2001 (05.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/245,564	6 November 2000 (06.11.2000)	US
60/245,565	6 November 2000 (06.11.2000)	US
60/245,572	6 November 2000 (06.11.2000)	US

(71) Applicant *(for all designated States except US)*: **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **SMOLYAR, Alex** [US/US]; 734 Boylston Street, Brookline, MA 02467 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

— *entirely in electronic form (except for this front page) and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/36613 A2

(54) Title: REGULATION OF HUMAN PATCHED-LIKE PROTEIN

(57) Abstract: Reagents which regulate human Patched-like protein and reagents which bind to human Patched-like protein gene products can play a role in preventing, ameliorating, or correcting dysfunctions or diseases including, but not limited to diabetes, cancer, cardiovascular diseases, and peripheral and central nervous system disorders. LIO 204-Foreign Countries

REGULATION OF HUMAN PATCHED-LIKE PROTEIN

This application incorporates by reference co-pending provisional applications Serial Nos. 60/245,572, 60/245,565 and 60/245,564 filed October 31, 2000.

TECHNICAL FIELD OF THE INVENTION

The invention relates to the area of regulation of human Patched-like protein.

BACKGROUND OF THE INVENTION

The 12 transmembrane domain protein Patched (PTCH) is the receptor for Sonic Hedgehog (Shh), a secreted molecule implicated in the formation of embryonic structures and in tumorigenesis. Binding of Sonic (Shh) protein to Patched protein prevents the normal inhibition of the G-protein coupled like receptor Smoothened (SMO) by PTCH (see Fig. 23). Hedgehog proteins, a family of secreted molecules first identified by a genetic screen in Drosophila, are involved in many patterning processes during development. Three mammalian hedgehog homologues have been identified: Sonic (Shh), Desert (Dhh), and Indian (Ihh). Shh acts to establish cell fate in the developing limb, somites, and neural tube. Ihh is involved specifically in chondrocyte development, and Dhh plays a key role in germ cell development. With the exception of the gut, in which both Ihh and Shh are expressed, the expression patterns of the hedgehog family members do not overlap. At the cell surface, Shh function appears to be mediated by a multicomponent receptor complex involving Patched (PTCH) and Smoothened(SMO), two multitransmembrane proteins initially identified as segment polarity genes in Drosophila and later characterized in vertebrates. Both genetic and biochemical evidence supports the existence of a receptor complex in which PTCH is the ligand-binding subunit and SMO, a G protein-coupled receptor-like molecule, is the signaling component. Upon binding of

Shh to PTCH, the normal inhibitory effect of PTCH on SMO is relieved, allowing SMO to transduce the Shh signal across the plasma membrane. In vertebrates a hedgehog gene family is involved in the control of left-right asymmetry, polarity in the central nervous system (CNS), somites and limb, organogenesis,chondrogenesis and spermatogenesis.

Loss of function mutations in the *Ptch* gene have been identified in patients with basal cell nevus syndrome, a hereditary disease characterized by multiple basal cell carcinomas. *See* U.S. Patent No. 6,027,882. Developmental abnormalities found with this syndrome include rib and craniofacial abnormalities, polydactyly, syndactyly and spina bifida. Tumors found with the syndrome include basal cell carcinomas, fibromas of the ovaries and heart, cysts of the skin, jaws and mesentery, meningiomas and medulloblastomas. *See* Gorlin, *Medicine 66*, 98-113 (1987).

Dysfunctional *Ptch* mutations also have been associated with a large percentage of sporadic basal cell carcinoma tumors. Other human cancers, *e.g.* basal cell carcinoma, transitional cell carcinoma of the bladder, meningiomas, medulloblastomas, etc., show decreased Patch activity, resulting from oncogenic mutations at the *Ptch* locus. Decreased Patch protein activity is also associated with inherited developmental abnormalities, *e.g.* rib and craniofacial abnormalities, polydactyly, syndactyly and spina bifida. *See* U.S. Patent No. 6,022,708.

Thus, there is a need in the art to identify novel human Patch-like genes which can be regulated and provide therapeutic options.

SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating a human Patched-like protein. This and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention is a patched-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 8;

the amino acid sequence shown in SEQ ID NO: 8;

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 14; and

the amino acid sequence shown in SEQ ID NO: 14.

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a patched-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 8;

the amino acid sequence shown in SEQ ID NO: 8;

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 14; and

the amino acid sequence shown in SEQ ID NO: 14.

Binding between the test compound and the patched-like protein polypeptide is detected. A test compound which binds to the patched-like protein polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the patched-like protein.

Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a patched-like protein polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 7;

the nucleotide sequence shown in SEQ ID NO: 7;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 13; and

the nucleotide sequence shown in SEQ ID NO: 13.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the patched-like protein through interacting with the patched-like protein mRNA.

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a patched-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 8;

the amino acid sequence shown in SEQ ID NO: 8;

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 14; and

the amino acid sequence shown in SEQ ID NO: 14.

A patched-like protein activity of the polypeptide is detected. A test compound which increases patched-like protein activity of the polypeptide relative to patched-

like protein activity in the absence of the test compound is thereby identified as a potential agent for increasing extracellular matrix degradation. A test compound which decreases patched-like protein activity of the polypeptide relative to patched-like protein activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a patched-like protein product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 7;

the nucleotide sequence shown in SEQ ID NO: 7;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 13; and

the nucleotide sequence shown in SEQ ID NO: 13.

Binding of the test compound to the patched-like protein product is detected. A test compound which binds to the patched-like protein product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a patched-like protein polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 7;

the nucleotide sequence shown in SEQ ID NO: 7;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 13; and

the nucleotide sequence shown in SEQ ID NO: 13.

Patched-like protein activity in the cell is thereby decreased.

The invention thus provides a human Patched-like protein which can be used to identify test compounds which may act, for example, as enhancers or inhibitors of formation of the receptor complex. Human Patched-like protein and fragments thereof also are useful in raising specific antibodies which can block the protein and effectively reduce its activity.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the DNA-sequence encoding a Patched-like protein polypeptide (SEQ ID NO:7).

Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO:8).

Fig. 3 shows the amino acid sequence of a protein from *Drosophila melanogaster* identified, Accession No. AE003784 (SEQ ID NO:9), which is homologues to Patched-like protein (SEQ ID NO:8).

Fig. 4 shows the DNA-sequence encoding a Patched-like protein Polypeptide (SEQ ID NO:10).

Fig. 5 shows the DNA-sequence encoding a Patched-like protein Polypeptide (SEQ ID NO:11).

Fig. 6 shows the DNA-sequence encoding a Patched-like protein Polypeptide (SEQ ID NO:12).

Fig. 7 shows the BLASTP alignment of human Patched-like protein (SEQ ID NO:7) with the protein from *Drosophila melanogaster*, Accession No. AE003784 (SEQ ID NO:9).

Fig. 8 shows the HMMPFAM - alignment of human Patched-like protein (SEQ ID NO:8) against pfam|hmm|Patched.

Fig. 9 shows the DNA-sequence encoding a patched-like protein polypeptide (SEQ ID NO:13).

Fig. 10 shows the amino acid sequence deduced from the DNA-sequence of Fig.9 (SEQ ID NO:14).

Fig. 11 shows the DNA-sequence encoding a patched-like protein polypeptide (SEQ ID NO:15).

Fig. 12 shows the DNA-sequence encoding a patched-like protein polypeptide (SEQ ID NO:16).

Fig. 13 shows the BLASTP alignment of human Patched-like protein (SEQ ID NO:13) with the protein from *Drosophila melanogaster*, Accession No. AE003784 (SEQ ID NO:9).

Fig. 14 shows the HMMPFAM - alignment of human Patched-like protein (SEQ ID NO:14) against pfam hmm Patched.

Fig. 15 shows the DNA-sequence encoding a Patched-like protein polypeptide (SEQ ID NO:1).

Fig. 16 shows the amino acid sequence deduced from the DNA-sequence of Fig.15 (SEQ ID NO:2).

Fig. 17 shows the amino acid sequence of a protein from *Caenorhabditis elegans* identified, Accession No. U88308 (SEQ ID NO:3), which is homologues to Patched-like protein (SEQ ID NO:2).

Fig. 18 shows the DNA-sequence encoding a Patched-like protein polypeptide (SEQ ID NO:4).

Fig. 19 shows the DNA-sequence encoding a Patched-like protein polypeptide (SEQ ID NO:5).

Fig. 20 shows the DNA-sequence encoding a Patched-like protein polypeptide (SEQ ID NO:6).

Fig. 21 shows the BLASTP alignment of human Patched-like protein (SEQ ID NO:13) with the protein from *Caenorhabditis elegans* identified, Accession No. U88308 (SEQ ID NO:3).

Fig. 22 shows the HMMPFAM - alignment of human Patched-like protein (SEQ ID NO:2) against pfam|hmm|Patched.

Fig. 23 illustrates schematically the signalling pathway "Patched" protein is involved in.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide encoding a patched-like protein polypeptide and being selected from the group consisting of:

a) a polynucleotide encoding a patched-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;
amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 8;
the amino acid sequence shown in SEQ ID NO: 8;
amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 14; and
the amino acid sequence shown in SEQ ID NO: 14.

b) a polynucleotide comprising the sequence of SEQ ID NO: 1, 7 or 13;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d).

Furthermore, it has been discovered by the present applicant that a novel Patched-like protein, particularly a human patched-like protein, is a discovery of the present invention. Human Patched-like protein comprises the amino acid sequence shown in SEQ ID NO:2, 8 or 14. A coding sequence for human Patched-like protein is shown

in SEQ ID NOS: 1, 7 and 13. Related ESTs are expressed in in mouse testis, fetal lung, and liver (SEQ ID NOS: 4-6), in adenocarcinoma and mouse cerebellum (SEQ ID NOS: 10-12) and in in ovary and mouse cerebellum (SEQ ID NOS: 15 and 16).

Human Patched-like protein is 26% identical over 525 amino acids to the *Drosophila melanogaster* protein identified with Accession No. AE003784 (Fig. 7). Human Patched-like protein of the invention is expected to be useful for the same purposes as previously identified human Patched proteins. Human Patched-like protein is believed to be useful in therapeutic methods to treat disorders such as diabetes, cancer, cardiovascular diseases, and peripheral and central nervous system disorders. Human Patched-like protein also can be used to screen for human Patched-like protein activators and inhibitors.

Polypeptides

Human Patched-like protein polypeptides according to the invention comprise at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 760, 770, 780, 790, or 792 contiguous amino acids selected from the amino acid sequence shown in SEQ ID NO: 2, 8 or 14 or a biologically active variant thereof, as defined below. A human Patched-like protein polypeptide of the invention therefore can be a portion of a human Patched-like protein, a full-length human Patched-like protein, or a fusion protein comprising all or a portion of a human Patched-like protein.

Biologically Active Variants

Human Patched-like protein polypeptide variants which are biologically active, *e.g.*, retain a Shh-binding activity, also are human Patched-like protein polypeptides. Preferably, naturally or non-naturally occurring human Patched-like protein polypeptide variants have amino acid sequences which are at least about 20, 25, 30, 35,

40, 45, 50, 55, 60, 65, or 70, preferably about 75, 80, 85, 90, 96, 96, or 98% identical to the amino acid sequence shown in SEQ ID NO: 2, 8 or 14 or a fragment thereof. Percent identity between a putative human Patched-like protein polypeptide variant and an amino acid sequence of SEQ ID NO: 2, 8 or 14 is determined by conventional methods. See, forexample, Altschul et al., Bull. Math. Bio. 48:603 (1986), and Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1992). Briefly, two amino acidsequences are aligned to optimize the alignment scores using a gap openingpenalty of 10, agap extension penalty of 1, and the "BLOSUM62" scoring matrix of Henikoffand Henikoff (ibid.).Those skilled in the art appreciate that there are many establishedalgorithms available to align two amino acid sequences. The "FASTA"similarity searchalgorithm of Pearson and Lipman is a suitable protein alignment method forexamining the level of identity shared by an amino acid sequence disclosedherein and theamino acid sequence of a putative variant. The FASTA algorithm is describedby Pearson and Lipman, Proc. Nat'l Acad. Sci. USA 85:2444(1988), and by Pearson, Meth. Enzymol. 183:63 (1990).Briefly, FASTA first characterizes sequence similarity by identifyingregions shared by the query sequence (e.g. SEQ ID NO: 2, 8 or 14) and a test sequencethat haveeither the highest density of identities (if the ktup variable is 1) orpairs of identities (if ktup=2), without considering conservative amino acidsubstitutions, insertions,or deletions. The ten regions with the highest density of identities are then rescored by comparing the similarity of all paired amino acids using anamino acidsubstitution matrix, and the ends of the regions are "trimmed" to includeonly those residues that contribute to the highest score. If there areseveral regions withscores greater than the "cutoff" value (calculated by a predeterminedformula based upon the length of the sequence and the ktup value), then thetrimmedinitialregions are examined to determine whether the regions can be joined to forman approximate alignment with gaps. Finally, the highest scoring regions ofthe twoamino acid sequences are aligned using a modification of theNeedleman-Wunsch- Sellers algorithm (Needleman and Wunsch, J. Mol. Biol.48:444 (1970); Sellers,SIAM J. Appl. Math. 26:787 (1974)), which allows for amino acid insertionsand deletions. Preferred parameters for FASTA

analysis are: ktup=1, gapopeningpenalty=10, gap extension penalty=1, and substitution matrix=BLOSUM62. These parameters can be introduced into a FASTA program by modifying thescoring matrix file ("SMATRIX"), as explained in Appendix 2 of Pearson,Meth. Enzymol. 183:63 (1990).FASTA can also be used to determine the sequence identity of nucleic acidmolecules using a ratio as disclosed above. For nucleotide sequencecomparisons,the ktup value can range between one to six, preferably from three to six,most preferably three, with other parameters set as default.

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a human Patched-like protein polypeptide can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active human Patched-like protein polypeptide can readily be determined by assaying for Shh-binding activity, as described for example, in Carpenter, *et al.*, *PROC. NATL. ACAD. SCI. U.S.A. 95*, 13630-34 (1998).

Fusion Proteins

Fusion proteins are useful for generating antibodies against human Patched-like protein polypeptide amino acid sequences and for use in various assay systems. For

example, fusion proteins can be used to identify proteins which interact with portions of a human Patched-like protein polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A human Patched-like protein polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 760, 770, 780 or 785 contiguous amino acids of SEQ ID NO:2 or at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 760, 770, 780, 790, or 792 contiguous amino acids of SEQ ID NO:8 or at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 530, or 536 contiguous amino acids of SEQ ID NO:14 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length human Patched-like protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between

the human Patched-like protein polypeptide-encoding sequence and the heterologous protein sequence, so that the human Patched-like protein polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from the complement of SEQ ID NO: 1, 7 or 13 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologs

Species homologs of human Patched-like protein polypeptide can be obtained using human Patched-like protein polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of human Patched-like protein polypeptide, and expressing the cDNAs as is known in the art.

Polynucleotides

A human Patched-like protein polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a human

Patched-like protein polypeptide. A coding sequence for human Patched-like protein is shown in SEQ ID NOS: 1, 7 and 13. Degenerate nucleotide sequences encoding human Patched-like protein polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to the nucleotide sequence shown in SEQ ID NO: 1, 7 or 13 or its complement also are human Patched-like protein polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of human Patched-like protein polynucleotides which encode biologically active human Patched-like protein polypeptides also are human Patched-like protein polynucleotides. Fragments comprising 8, 10, 25, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, or 2355 contiguous nucleotides of SEQ ID NO:1 or 8, 10, 25, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, or 2379 contiguous nucleotides of SEQ ID NO:7 or 8, 10, 25, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, or 1650 contiguous nucleotides of SEQ ID NO:13 or its complement also are human Patched-like protein polynucleotides.

Identification of Polynucleotide Variants and Homologs

Variants and homologs of the human Patched-like protein polynucleotides described above also are human Patched-like protein polynucleotides. Typically, homologous human Patched-like protein polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known human Patched-like protein polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1%

SDS, 50°C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each--homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the human Patched-like protein polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of human Patched-like protein polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5°C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol. 81*, 123 (1973). Variants of human Patched-like protein polynucleotides or human Patched-like protein polynucleotides of other species can therefore be identified by hybridizing a putative homologous human Patched-like protein polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1, 7 or 13 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to human Patched-like protein polynucleotides or their complements following stringent hybridization and/or wash conditions also are human Patched-like protein polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20°C below the

calculated T_m of the hybrid under study. The T_m of a hybrid between a human Patched-like protein polynucleotide having a nucleotide sequence shown in SEQ ID NO: 1, 7 or 13 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A. 48*, 1390 (1962):

$$T_m = 81.5\ ^\circ C - 16.6(\log_{10}[Na^+]) + 0.41(\%G + C) - 0.63(\%\text{formamide}) - 600/l),$$

where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65°C, or 50% formamide, 4X SSC at 42°C, or 0.5X SSC, 0.1% SDS at 65°C. Highly stringent wash conditions include, for example, 0.2X SSC at 65°C.

Preparation of Polynucleotides

A human Patched-like protein polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated human Patched-like protein polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments which comprises Patched-like nucleotide sequences. Isolated polynucleotides are in preparations which are free or at least 70, 80, or 90% free of other molecules.

Human Patched-like protein cDNA molecules can be made with standard molecular biology techniques, using human Patched-like protein mRNA as a template. Human

Patched-like protein cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesizes human Patched-like protein polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a human Patched-like protein polypeptide having, for example, an amino acid sequence shown in SEQ ID NO:2, 8 or 14 or a biologically active variant thereof.

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic. 2*, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res. 16*, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction

enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template. Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al., PCR Methods Applic. 1*, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al., Nucleic Acids Res. 19*, 3055-3060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data

display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

Obtaining Polypeptides

Human Patched-like protein polypeptides can be obtained, for example, by purification from human cells, by expression of human Patched-like protein polynucleotides, or by direct chemical synthesis.

Protein Purification

Human Patched-like protein polypeptides can be purified from any cell which expresses the enzyme, including host cells which have been transfected with human Patched-like protein expression constructs. A purified human Patched-like protein polypeptide is separated from other compounds which normally associate with the human Patched-like protein polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified human Patched-like protein polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

Expression of Polynucleotides

To express a human Patched-like protein polynucleotide, the polynucleotide can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are

well known to those skilled in the art can be used to construct expression vectors containing sequences encoding human Patched-like protein polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a human Patched-like protein polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian

viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a human Patched-like protein polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the human Patched-like protein polypeptide. For example, when a large quantity of a human Patched-like protein polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the human Patched-like protein polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem. 264*, 5503-5509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol. 153*, 516-544, 1987.

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding human Patched-like protein polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV (Takamatsu, *EMBO J. 6*, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J. 3*, 1671-1680, 1984; Broglie *et al.*, *Science 224*, 838-843, 1984; Winter *et al.*, *Results Probl. Cell Differ. 17*, 85-105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (*e.g.*, Hobbs or Murray, in MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

An insect system also can be used to express a human Patched-like protein polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding human Patched-like protein polypeptides can be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of human Patched-like protein polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect *S. frugiperda cells* or *Trichoplusia* larvae in which human Patched-like protein polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci. 91*, 3224-3227, 1994).

Mammalian Expression Systems

A number of viral-based expression systems can be used to express human Patched-like protein polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding human Patched-like protein polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing a human Patched-like protein polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci. 81*, 3655-3659, 1984). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding human Patched-like protein polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a human Patched-like protein polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell

system which is used (see Scharf *et al.*, *Results Probl. Cell Differ. 20*, 125-162, 1994).

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed human Patched-like protein polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express human Patched-like protein polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced human Patched-like protein sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines.

These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell 11*, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell 22*, 817-23, 1980) genes which can be employed in *tk*⁻ or *aprt*⁻ cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci. 77*, 3567-70, 1980), *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol. 150*, 1-14, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. 85*, 8047-51, 1988). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol. 55*, 121-131, 1995).

Detecting Expression

Although the presence of marker gene expression suggests that the human Patched-like protein polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a human Patched-like protein polypeptide is inserted within a marker gene sequence, transformed cells containing sequences which encode a human Patched-like protein polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a human Patched-like protein polypeptide under the control of a single promoter. Expression of the marker gene in response to

induction or selection usually indicates expression of the human Patched-like protein polynucleotide.

Alternatively, host cells which contain a human Patched-like protein polynucleotide and which express a human Patched-like protein polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding a human Patched-like protein polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a human Patched-like protein polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a human Patched-like protein polypeptide to detect transformants which contain a human Patched-like protein polynucleotide.

A variety of protocols for detecting and measuring the expression of a human Patched-like protein polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a human Patched-like protein polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med. 158*, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for

producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding human Patched-like protein polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a human Patched-like protein polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a human Patched-like protein polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode human Patched-like protein polypeptides can be designed to contain signal sequences which direct secretion of soluble human Patched-like protein polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound human Patched-like protein polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a human Patched-like protein polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such

purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the human Patched-like protein polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a human Patched-like protein polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in Porath *et al.*, *Prot. Exp. Purif. 3*, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the human Patched-like protein polypeptide from the fusion protein. Vectors which contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol. 12*, 441-453, 1993.

Chemical Synthesis

Sequences encoding a human Patched-like protein polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al. Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a human Patched-like protein polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc. 85*, 2149-2154, 1963; Roberge *et al.*, *Science 269*, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of human Patched-like protein polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The composition of a synthetic human Patched-like protein polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the human Patched-like protein polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce human Patched-like protein polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter human Patched-like protein polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a human Patched-like protein polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv, which are capable of binding an epitope of a human Patched-like protein polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a human Patched-like protein polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to a human Patched-like protein polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to human Patched-like polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a human Patched-like protein polypeptide from solution.

Human Patched-like protein polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal

antibodies. If desired, a human Patched-like protein polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (*e.g.*, aluminum hydroxide), and surface active substances (*e.g.* lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies which specifically bind to a human Patched-like protein polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature 256*, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods 81*, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci. 80*, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol. 62*, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci. 81*, 6851-6855, 1984; Neuberger *et al.*, *Nature 312*, 604-608, 1984; Takeda *et al.*, *Nature 314*, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in

GB2188638B. Antibodies which specifically bind to a human Patched-like protein polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies which specifically bind to human Patched-like protein polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton, *Proc. Natl. Acad. Sci. 88*, 11120-23, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev. 5*, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol. 15*, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem. 269*, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer 61*, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth. 165*, 81-91).

Antibodies which specifically bind to human Patched-like protein polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as

disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci. 86*, 3833-3837, 1989; Winter *et al.*, *Nature 349*, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a human Patched-like protein polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of human Patched-like protein gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such

alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. *See* Brown, *Meth. Mol. Biol. 20*, 1-8, 1994; Sonveaux, *Meth. Mol. Biol. 26*, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev. 90*, 543-583, 1990.

Modifications of human Patched-like protein gene expression can be obtained by designing antisense oligonucleotides which will form duplexes to the control, 5', or regulatory regions of the human Patched-like protein gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a human Patched-like protein polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a human Patched-like protein polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent human Patched-like protein nucleotides, can provide sufficient targeting specificity for human Patched-like protein mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an

antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular human Patched-like protein polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a human Patched-like protein polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. *See, e.g.*, Agrawal *et al.*, *Trends Biotechnol. 10*, 152-158, 1992; Uhlmann *et al.*, *Chem. Rev. 90*, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett. 215*, 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. *See, e.g.*, Cech, *Science 236*, 1532-1539; 1987; Cech, *Ann. Rev. Biochem. 59*, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol. 2*, 605-609; 1992, Couture & Stinchcomb, *Trends Genet. 12*, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.*, Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a human Patched-like protein polynucleotide can be used to generate ribozymes which will specifically bind to mRNA transcribed from the human Patched-like protein polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (*see* Haseloff *et al. Nature 334*, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a human Patched-like protein RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate human Patched-like protein RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease human Patched-like protein expression. Alternatively, if it is desired that the cells stably retain the DNA

construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors which induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Differentially Expressed Genes

Described herein are methods for the identification of genes whose products interact with human Patched-like protein. Such genes may represent genes which are differentially expressed in disorders including, but not limited to diabetes, cancer, cardiovascular diseases, and peripheral and central nervous system disorders. Further, such genes may represent genes which are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human Patched-like protein gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by

which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

Identification of Differentially Expressed Genes

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique which does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed.,, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples which represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature 308*, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 2825, 1984), and, preferably, differential display (Liang & Pardee, *Science 257*, 967-71, 1992; U.S. Patent 5,262,311).

The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human Patched-like protein. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human Patched-like protein. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human Patched-like protein gene or gene product are up-regulated or down-regulated.

Screening Methods

The invention provides assays for screening test compounds which bind to or modulate the activity of a human Patched-like protein polypeptide or a human Patched-like protein polynucleotide. A test compound preferably binds to a human Patched-like protein polypeptide or polynucleotide. More preferably, a test compound decreases or increases human Patched-like protein activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. *See* Lam, *Anticancer Drug Des. 12*, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (*see*, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6909, 1993; Erb *et al. Proc. Natl. Acad. Sci. U.S.A. 91*, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem. 37*, 2678, 1994; Cho *et al.*, *Science 261*, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.*

33, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2061; Gallop *et al.*, *J. Med. Chem. 37*, 1233, 1994). Libraries of compounds can be presented in solution (*see, e.g.*, Houghten, *BioTechniques 13*, 412-421, 1992), or on beads (Lam, *Nature 354*, 82-84, 1991), chips (Fodor, *Nature 364*, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 89*, 1865-1869, 1992), or phage (Scott & Smith, *Science 249*, 386-390, 1990; Devlin, *Science 249*, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci. 97*, 6378-6382, 1990; Felici, *J. Mol. Biol. 222*, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

Test compounds can be screened for the ability to bind to human Patched-like protein polypeptides or polynucleotides or to affect human Patched-like protein activity or human Patched-like protein gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 19*, 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity 2*, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

For binding assays, the test compound is preferably a small molecule which binds to and occupies, for example, the active site of the human Patched-like protein polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or the human Patched-like protein polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound which is bound to the human Patched-like protein polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a human Patched-like protein polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a human Patched-like protein polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a human Patched-like protein polypeptide (McConnell *et al.*, *Science 257*, 1906-1912, 1992).

Determining the ability of a test compound to bind to a human Patched-like protein polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem. 63*, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol. 5*, 699-705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a human Patched-like protein polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell 72*, 223-232, 1993; Madura *et al.*, *J. Biol. Chem.*

268, 12046-12054, 1993; Bartel *et al.*, *BioTechniques 14*, 920-924, 1993; Iwabuchi *et al.*, *Oncogene 8*, 1693-1696, 1993; and Brent W094/10300), to identify other proteins which bind to or interact with the human Patched-like protein polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide encoding a human Patched-like protein polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the human Patched-like protein polypeptide.

It may be desirable to immobilize either the human Patched-like protein polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the human Patched-like protein polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be

used to attach the enzyme polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a human Patched-like protein polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the human Patched-like protein polypeptide is a fusion protein comprising a domain that allows the human Patched-like protein polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed human Patched-like protein polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a human Patched-like protein polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated human Patched-like protein polypeptides (or polynucleotides) or test compounds can be prepared from biotin-NHS(N-hydroxysuccinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in

the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a human Patched-like protein polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the active site of the human Patched-like protein polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the human Patched-like protein polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the human Patched-like protein polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a human Patched-like protein polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a human Patched-like protein polypeptide or polynucleotide can be used in a cell-based assay system. A human Patched-like protein polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a human Patched-like protein polypeptide or polynucleotide is determined as described above.

Gene Expression

In another embodiment, test compounds which increase or decrease human Patched-like protein gene expression are identified. A human Patched-like protein polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the human Patched-like protein polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in

the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of human Patched-like protein mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a human Patched-like protein polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a human Patched-like protein polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses a human Patched-like protein polynucleotide can be used in a cell-based assay system. The human Patched-like protein polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions which can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a human Patched-like protein polypeptide,

human Patched-like protein polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a human Patched-like protein polypeptide, or mimetics, activators, or inhibitors of a human Patched-like protein polypeptide activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxy-propylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating

or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular

barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic Indications and Methods

Human patch-like protein may be regulated to treat diabetes, cancer, cardiovascular diseases, and peripheral and central nervous system disorders.

Cancer

Cancer is a disease fundamentally caused by oncogenic cellular transformation. There are several hallmarks of transformed cells that distinguish them from their

normal counterparts and underlie the pathophysiology of cancer. These include uncontrolled cellular proliferation, unresponsiveness to normal death-inducing signals (immortalization), increased cellular motility and invasiveness, increased ability to recruit blood supply through induction of new blood vessel formation (angiogenesis), genetic instability, and dysregulated gene expression. Various combinations of these aberrant physiologies, along with the acquisition of drug-resistance frequently lead to an intractable disease state in which organ failure and patient death ultimately ensue.

Most standard cancer therapies target cellular proliferation and rely on the differential proliferative capacities between transformed and normal cells for their efficacy. This approach is hindered by the facts that several important normal cell types are also highly proliferative and that cancer cells frequently become resistant to these agents. Thus, the therapeutic indices for traditional anti-cancer therapies rarely exceed 2.0.

The advent of genomics-driven molecular target identification has opened up the possibility of identifying new cancer-specific targets for therapeutic intervention that will provide safer, more effective treatments for cancer patients. Thus, newly discovered tumor-associated genes and their products can be tested for their role(s) in disease and used as tools to discover and develop innovative therapies. Genes playing important roles in any of the physiological processes outlined above can be characterized as cancer targets.

Genes or gene fragments identified through genomics can readily be expressed in one or more heterologous expression systems to produce functional recombinant proteins. These proteins are characterized in vitro for their biochemical properties and then used as tools in high-throughput molecular screening programs to identify chemical modulators of their biochemical activities. Agonists and/or antagonists of target protein activity can be identified in this manner and subsequently tested in cellular

and in vivo disease models for anti-cancer activity. Optimization of lead compounds with iterative testing in biological models and detailed pharmacokinetic and toxicological analyses form the basis for drug development and subsequent testing in humans.

Cardiovascular diseases

Cardiovascular diseases include the following disorders of the heart and the vascular system: congestive heart failure, myocardial infarction, ischemic diseases of the heart, all kinds of atrial and ventricular arrhythmias, hypertensive vascular diseases, and peripheral vascular diseases.

Heart failure is defined as a pathophysiologic state in which an abnormality of cardiac function is responsible for the failure of the heart to pump blood at a rate commensurate with the requirement of the metabolizing tissue. It includes all forms of pumping failure, such as high-output and low-output, acute and chronic, right-sided or left-sided, systolic or diastolic, independent of the underlying cause.

Myocardial infarction (MI) is generally caused by an abrupt decrease in coronary blood flow that follows a thrombotic occlusion of a coronary artery previously narrowed by arteriosclerosis. MI prophylaxis (primary and secondary prevention) is included, as well as the acute treatment of MI and the prevention of complications.

Ischemic diseases are conditions in which the coronary flow is restricted resulting in a perfusion which is inadequate to meet the myocardial requirement for oxygen. This group of diseases includes stable angina, unstable angina, and asymptomatic ischemia.

Arrhythmias include all forms of atrial and ventricular tachyarrhythmias (atrial tachycardia, atrial flutter, atrial fibrillation, atrio-ventricular reentrant tachycardia,

preexcitation syndrome, ventricular tachycardia, ventricular flutter, and ventricular fibrillation), as well as bradycardic forms of arrhythmias.

Hypertensive vascular diseases include primary as well as all kinds of secondary arterial hypertension (renal, endocrine, neurogenic, others). The disclosed gene and its product may be used as drug targets for the treatment of hypertension as well as for the prevention of all complications.

Peripheral vascular diseases are defined as vascular diseases in which arterial and/or venous flow is reduced resulting in an imbalance between blood supply and tissue oxygen demand. It includes chronic peripheral arterial occlusive disease (PAOD), acute arterial thrombosis and embolism, inflammatory vascular disorders, Raynaud's phenomenon, and venous disorders.

Peripheral or central nervous system disorders

Peripheral and central nervous system disorders which may be treated include brain injuries, cerebrovascular diseases and their consequences, Parkinson's disease, corticobasal degeneration, motor neuron disease, dementia, including ALS, multiple sclerosis, traumatic brain injury, stroke, post-stroke, post-traumatic brain injury, and small-vessel cerebrovascular disease. Dementias, such as Alzheimer's disease, vascular dementia, dementia with Lewy bodies, frontotemporal dementia and Parkinsonism linked to chromosome 17, frontotemporal dementias, including Pick's disease, progressive nuclear palsy, corticobasal degeneration, Huntington's disease, thalamic degeneration, Creutzfeld-Jakob dementia, HIV dementia, schizophrenia with dementia, and Korsakoff's psychosis also can be treated. Similarly, it may be possible to treat cognitive-related disorders, such as mild cognitive impairment, age-associated memory impairment, age-related cognitive decline, vascular cognitive impairment, attention deficit disorders, attention deficit hyperactivity disorders, and

memory disturbances in children with learning disabilities, by regulating the activity of human Patched-like protein.

Pain associated with peripheral or central nervous system disorders also can be treated by regulating the activity of human Patched-like protein. Pain which can be treated includes that associated with central nervous system disorders, such as multiple sclerosis, spinal cord injury, sciatica, failed back surgery syndrome, traumatic brain injury, epilepsy, Parkinson's disease, post-stroke, and vascular lesions in the brain and spinal cord (*e.g.*, infarct, hemorrhage, vascular malformation). Non-central neuropathic pain includes that associated with post mastectomy pain, reflex sympathetic dystrophy (RSD), trigeminal neuralgiaradioculopathy, post-surgical pain, HIV/AIDS related pain, cancer pain, metabolic neuropathies (*e.g.*, diabetic neuropathy, vasculitic neuropathy secondary to connective tissue disease), paraneoplastic polyneuropathy associated, for example, with carcinoma of lung, or leukemia, or lymphoma, or carcinoma of prostate, colon or stomach, trigeminal neuralgia, cranial neuralgias, and post-herpetic neuralgia. Pain associated with cancer and cancer treatment also can be treated, as can headache pain (for example, migraine with aura, migraine without aura, and other migraine disorders), episodic and chronic tension-type headache, tension-type like headache, cluster headache, and chronic paroxysmal hemicrania.

Diabetes

Diabetes mellitus is a common metabolic disorder characterized by an abnormal elevation in blood glucose, alterations in lipids and abnormalities (complications) in the cardiovascular system, eye, kidney and nervous system. Diabetes is divided into two separate diseases: type 1 diabetes (juvenile onset), which results from a loss of cells which make and secrete insulin, and type 2 diabetes (adult onset), which is caused by a defect in insulin secretion and a defect in insulin action.

Type 1 diabetes is initiated by an autoimuune reaction that attacks the insulin secreting cells (beta cells) in the pancreatic islets. Agents that prevent this reaction from occurring or that stop the reaction before destruction of the beta cells has been accomplished are potential therapies for this disease. Other agents that induce beta cell proliferation and regeneration also are potential therapies.

Type II diabetes is the most common of the two diabetic conditions (6% of the population). The defect in insulin secretion is an important cause of the diabetic condition and results from an inability of the beta cell to properly detect and respond to rises in blood glucose levels with insulin release. Therapies that increase the response by the beta cell to glucose would offer an important new treatment for this disease.

The defect in insulin action in Type II diabetic subjects is another target for therapeutic intervention. Agents that increase the activity of the insulin receptor in muscle, liver, and fat will cause a decrease in blood glucose and a normalization of plasma lipids. The receptor activity can be increased by agents that directly stimulate the receptor or that increase the intracellular signals from the receptor. Other therapies can directly activate the cellular end process, *i.e.* glucose transport or various enzyme systems, to generate an insulin-like effect and therefore a produce beneficial outcome. Because overweight subjects have a greater susceptibility to Type II diabetes, any agent that reduces body weight is a possible therapy.

Both Type I and Type diabetes can be treated with agents that mimic insulin action or that treat diabetic complications by reducing blood glucose levels. Likewise, agents that reduces new blood vessel growth can be used to treat the eye complications that develop in both diseases.

This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a

test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a human Patched-like protein polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

A reagent which affects human Patched-like protein activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce human Patched-like protein activity. The reagent preferably binds to an expression product of a human Patched-like protein gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells which have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its

contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to poly-ethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods which are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al. Trends in Biotechnol. 11*, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem. 263*, 621-24 (1988);

Wu *et al., J. Biol. Chem. 269*, 542-46 (1994); Zenke *et al., Proc. Natl. Acad. Sci. U.S.A. 87*, 3655-59 (1990); Wu *et al., J. Biol. Chem. 266*, 338-42 (1991).

Determination of a Therapeutically Effective Dose

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases human Patched-like protein activity relative to the human Patched-like protein activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

Therapeutic efficacy and toxicity, *e.g.*, ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD_{50}/ED_{50}.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors which can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 µg to about 50 µg/kg, about 50 µg to about 5 mg/kg, about 100 µg to about 500 µg/kg of patient body weight, and about 200 to about 250 µg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo*

dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 μg to about 2 mg, about 5 μg to about 500 μg, and about 20 μg to about 100 μg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides which express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a human Patched-like protein gene or the activity of a human Patched-like protein polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a human Patched-like protein gene or the activity of a human Patched-like protein polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to human Patched-like protein-specific mRNA, quantitative RT-PCR, immunologic detection of a human Patched-like protein polypeptide, or measurement of human Patched-like protein activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

Diagnostic Methods

Human Patched-like protein also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences which encode the enzyme. For example, differences can be determined between the cDNA or genomic sequence encoding human Patched-like protein in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for example, by high resolution gel electrophoresis. DNA fragments of different sequences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions

according to their specific melting or partial melting temperatures (*see, e.g.*, Myers *et al.*, *Science 230*, 1242, 1985). Sequence changes at specific locations can also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g.*, Cotton *et al.*, *Proc. Natl. Acad. Sci. USA 85*, 4397-4401, 1985). Thus, the detection of a specific DNA sequence can be performed by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes and Southern blotting of genomic DNA. In addition to direct methods such as gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Altered levels of a human Patched-like protein also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1

Detection of patched-like protein activity

The polynucleotide of SEQ ID NO: 1, 7 or 13 is inserted into the expression vector pCEV4 and the expression vector pCEV4-patched-like protein polypeptide obtained is transfected into human embryonic kidney 293 cells. From these cells extracts are obtained and centrifuged at 1000 rpm for 5 minutes at 4°C. The supernatant is centrifuged at 30,000 x g for 20 minutes at 4°C. The pellet is suspended in binding buffer containing 50 mM Tris HCl, 5 mM $MgSO_4$, 1 mM EDTA, 100 mM NaCl, pH 7.5, supplemented with 0.1 % BSA, 2 µg/ml aprotinin, 0.5 mg/ml leupeptin, and 10 µg/ml phosphoramidon. Optimal membrane suspension dilutions, defined as the protein concentration required to bind less than 10% of the added radioligand, i.e. sonic hadgehog (SHH), are added to 96-well polypropylene microtiter plates containing ^{125}I-labeled ligand or test compound, non-labeled peptides, and binding buffer to a final volume of 250 µl.

In equilibrium saturation binding assays, membrane preparations are incubated in the presence of increasing concentrations (0.1 nM to 4 nM) of ^{125}I-labeled ligand or test compound (specific activity 2200 Ci/mmol). The binding affinities of different test compounds are determined in equilibrium competition binding assays, using 0.1 nM ^{125}I-peptide in the presence of twelve different concentrations of each test compound.

Binding reaction mixtures are incubated for one hour at 30°C. The reaction is stopped by filtration through GF/B filters treated with 0.5% polyethyleneimine, using a cell harvester. Radioactivity is measured by scintillation counting, and data are analyzed by a computerized non-linear regression program.

Non-specific binding is defined as the amount of radioactivity remaining after incubation of membrane protein in the presence of 100 nM of unlabeled peptide. Protein concentration is measured by the Bradford method using Bio-Rad Reagent,

with bovine serum albumin as a standard. It is shown that the polypeptide of SEQ ID NO: 2, 8 and 14 respectively have a patched-like protein activity.

EXAMPLE 2

Expression of recombinant human Patched-like protein

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of recombinant human Patched-like polypeptides in yeast. The human Patched-like protein-encoding DNA sequence is derived from SEQ ID NO:1, 7 or 13. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5'-end an initiation codon and at its 3'-end an enterokinase cleavage site, a His6 reporter tag and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added and after digestion of the multiple cloning site of pPICZ B with the corresponding restriction enzymes the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is used to transform the yeast.

The yeast is cultivated under usual conditions in 5 liter shake flasks and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified human Patched-like protein polypeptide is obtained.

EXAMPLE 3

Identification of test compounds that bind to human Patched-like protein polypeptides

Purified human Patched-like protein polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Human Patched-like protein polypeptides comprise the amino acid sequence shown in SEQ ID NO:2, 8 or 14. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a human Patched-like protein polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound which increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a human Patched-like protein polypeptide.

EXAMPLE 4

Identification of a test compound which increases human Patched-like protein gene expression

A test compound is administered to a culture of human cells transfected with a human Patched-like protein expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells which have not been transfected is incubated for the same time without the test compound to provide a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem. 18*, 5294-99, 1979). Northern blots are prepared using 20 to 30 μg total RNA and

hybridized with a ^{32}P-labeled human Patched-like protein-specific probe at 65°C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NO:1, 7 or 13. A test compound which decreases the human Patched-like protein-specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of human Patched-like protein gene expression.

EXAMPLE 5

Identification of a test compound which increases human Patched-like protein activity

A test compound is administered to a culture of human cells co-transfected with a Flag-tagged human Patched-like protein expression construct and a SMO expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells which have not been transfected is incubated for the same time without the test compound to provide a negative control. Human Patched-like protein activity is measured using the method of Carpenter, *et al.*, *PROC. NATL. ACAD. SCI. U.S.A. 95*, 13630-34 (1998).

Briefly, Patched-SMO complexes are detected by coimmunoprecipitation from dual-transfected cells. Protein A-bound Flag-specific antibodies are used to immuno-precipitate the Patched-SMO complexes. Complexes are then separated on a 6% acrylamide gels, transferred to nitrocellulose, and detected using Flag-specific antibodies in conjunction with an enhanced chemiluminescent detection system (Amersham).

A test compound which increases the SMO-binding activity of the human Patched-like protein relative to the SMO-binding activity in the absence of the test compound is identified as an enhancer of human Patched-like protein activity.

EXAMPLE 6

Tissue-specific expression of human Patched-like protein

The qualitative expression pattern of human Patched-like protein in various tissues is determined by Reverse Transcription-Polymerase Chain Reaction (RT-PCR). To demonstrate that human Patched-like protein is involved in cancer, expression is determined in the following tissues: skin, ovary, brain, and cerebellum. Expression in the following cancer cell lines also is determined: DU-145 (prostate), NCI-H125 (lung), HT-29 (colon), COLO-205 (colon), A-549 (lung), NCI-H460 (lung), HT-116 (colon), DLD-1 (colon), MDA-MD-231 (breast), LS174T (colon), ZF-75 (breast), MDA-MN-435 (breast), HT-1080, MCF-7 (breast), and U87. Matched pairs of malignant and normal tissue from the same patient also are tested.

To demonstrate that human Patched-like protein is involved in cardiovascular disease, expression is determined in the following tissues: muscle, heart, lung, placenta, skin, and peripheral blood lymphocytes. As a final step, the expression of human Patched-like protein in cells derived from normal individuals is compared with the expression of cells derived from individuals with cardiovascular disorders.

To demonstrate that human Patched-like protein is involved in peripheral or central nervous system disorders, the following tissues are screened: fetal and adult brain, muscle, heart, lung, kidney, liver, thymus, testis, colon, placenta, trachea, pancreas, kidney, gastric mucosa, colon, liver, cerebellum, skin, cortex (Alzheimer's and normal), hypothalamus, cortex, amygdala, cerebellum, hippocampus, choroid, plexus, thalamus, and spinal cord. As a final step, the expression of human Patched-like protein in cells derived from normal individuals is compared with the expression of cells derived from individuals with peripheral or central nervous system disorders.

Quantitative expression profiling. Quantitative expression profiling is performed by the form of quantitative PCR analysis called "kinetic analysis" firstly described in

Higuchi *et al., BioTechnology 10*, 413-17, 1992, and Higuchi *et al., BioTechnology 11*, 1026-30, 1993. The principle is that at any given cycle within the exponential phase of PCR, the amount of product is proportional to the initial number of template copies.

If the amplification is performed in the presence of an internally quenched fluorescent oligonucleotide (TaqMan probe) complementary to the target sequence, the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA polymerase and a fluorescent dye released in the medium (Holland *et al., Proc. Natl. Acad. Sci. U.S.A. 88*, 7276-80, 1991). Because the fluorescence emission will increase in direct proportion to the amount of the specific amplified product, the exponential growth phase of PCR product can be detected and used to determine the initial template concentration (Heid *et al., Genome Res. 6*, 986-94, 1996, and Gibson *et al., Genome Res. 6*, 995-1001, 1996).

The amplification of an endogenous control can be performed to standardize the amount of sample RNA added to a reaction. In this kind of experiment, the control of choice is the 18S ribosomal RNA. Because reporter dyes with differing emission spectra are available, the target and the endogenous control can be independently quantified in the same tube if probes labeled with different dyes are used.

All "real time PCR" measurements of fluorescence are made in the ABI Prism 7700.

RNA extraction and cDNA preparation. The total RNAs used for expression quantification are listed below along with their suppliers, if commercially available. RNAs labeled "from autopsy" were extracted from autoptic tissues with the TRIzol reagent (Life Technologies, MD) according to the manufacturer's protocol.

Fifty µg of each RNA were treated with DNase I for 1 hour at 37°C in the following reaction mix: 0.2 U/µl RNase-free DNase I (Roche Diagnostics, Germany); 0.4 U/µl

RNase inhibitor (PE Applied Biosystems, CA); 10 mM Tris-HCl pH 7.9; 10 mM $MgCl_2$; 50 mM NaCl; and 1 mM DTT.

After incubation, RNA is extracted once with 1 volume of phenol:chloroform:isoamyl alcohol (24:24:1) and once with chloroform, and precipitated with 1/10 volume of 3 M NaAcetate, pH 5.2, and 2 volumes of ethanol.

Fifty µg of each RNA from the autoptic tissues are DNase treated with the DNA-free kit purchased from Ambion (Ambion, TX). After resuspension and spectrophotometric quantification, each sample is reverse transcribed with the TaqMan Reverse Transcription Reagents (PE Applied Biosystems, CA) according to the manufacturer's protocol. The final concentration of RNA in the reaction mix is 200 ng/µl. Reverse transcription is carried out with 2.5 µM of random hexamer primers.

TaqMan quantitative analysis. Specific primers and probe are designed according to the recommendations of PE Applied Biosystems and are listed below:

forward primer: 5'-(gene specific sequence)-3'
reverse primer: 5'-(gene specific sequence)-3'
probe: 5'-(FAM)-(gene specific sequence)-(TAMRA)-3'
where FAM = 6-carboxy-fluorescein
and TAMRA = 6-carboxy-tetramethyl-rhodamine.

The expected length of the PCR product is -(gene specific length)-bp.

Quantification experiments are performed on 10 ng of reverse transcribed RNA from each sample. Each determination is done in triplicate.

Total cDNA content is normalized with the simultaneous quantification (multiplex PCR) of the 18S ribosomal RNA using the Pre-Developed TaqMan Assay Reagents (PDAR) Control Kit (PE Applied Biosystems, CA).

The assay reaction mix is as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 1X PDAR control - 18S RNA (from 20X stock); 300 nM forward primer; 900 nM reverse primer; 200 nM probe; 10 ng cDNA; and water to 25 µl.

Each of the following steps are carried out once: pre PCR, 2 minutes at 50°C, and 10 minutes at 95°C. The following steps are carried out 40 times: denaturation, 15 seconds at 95°C, annealing/extension, 1 minute at 60°C.

The experiment is performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR are processed as described in the ABI Prism 7700 user's manual in order to achieve better background subtraction as well as signal linearity with the starting target quantity.

REFERENCES

1. Carpenter, *et al.*, *PROC. NATL. ACAD. SCI. U.S.A. 95(23)*, 13630-34 (1998).
2. Alcedo and Noll, *Biol. Chem. 378(7)*, 583-90 (1997).
3. McMahon, *Cell 100(2)*, 185-88 (2000).
4. Burke, *et al.*, *Cell 99(7)*, 803-15 (1999).
5. Goodrich and Scott, *Neuron 21(6)*, 1243-57 (1998).
6. Currie, *J. Mol. Med. 76(6)*, 421-33 (1998).

CLAIMS

1. An isolated polynucleotide encoding a patched-like protein polypeptide and being selected from the group consisting of:

a) a polynucleotide encoding a patched-like protein polypeptide comprising an amino acid sequence selected form the group consisting of:

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 8;

the amino acid sequence shown in SEQ ID NO: 8;

amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO: 14; and

the amino acid sequence shown in SEQ ID NO: 14.

b) a polynucleotide comprising the sequence of SEQ ID NO: 1, 7 or 13;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a to (d).

2. An expression vector containing any polynucleotide of claim 1.

3. A host cell containing the expression vector of claim 2.

4. A substantially purified patched-like protein polypeptide encoded by a polynucleotide of claim 1.

5. A method for producing a patched-like protein polypeptide, wherein the method comprises the following steps:

a) culturing the host cell of claim 3 under conditions suitable for the expression of the patched-like protein polypeptide; and

b) recovering the patched-like protein polypeptide from the host cell culture.

6. A method for detection of a polynucleotide encoding a patched-like protein polypeptide in a biological sample comprising the following steps:

a) hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and

b) detecting said hybridization complex.

7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

8. A method for the detection of a polynucleotide of claim 1 or a patched-like protein polypeptide of claim 4 comprising the steps of:

contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the patched-like protein polypeptide.

9. A diagnostic kit for conducting the method of any one of claims 6 to 8.

10. A method of screening for agents which decrease the activity of a patched-like protein, comprising the steps of:

contacting a test compound with any patched-like protein polypeptide encoded by any polynucleotide of claim1;

detecting binding of the test compound to the patched-like protein polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a patched-like protein.

11. A method of screening for agents which regulate the activity of a patched-like protein, comprising the steps of:

contacting a test compound with a patched-like protein polypeptide encoded by any polynucleotide of claim 1; and

detecting a patched-like protein activity of the polypeptide, wherein a test compound which increases the patched-like protein activity is identified as a potential therapeutic agent for increasing the activity of the patched-like protein, and wherein a test compound which decreases the patched-like protein activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the patched-like protein.

12. A method of screening for agents which decrease the activity of a patched-like protein, comprising the steps of:

contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of patched-like protein.

13. A method of reducing the activity of patched-like protein, comprising the steps of:

contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any patched-like protein polypeptide of claim 4, whereby the activity of patched-like protein is reduced.

14. A reagent that modulates the activity of a patched-like protein polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.

15. A pharmaceutical composition, comprising:

the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

16. Use of the expression vector of claim 2 or the reagent of claim 14 in the preparation of a medicament for modulating the activity of a patched-like protein in a disease.

17. Use of claim 16 wherein the disease is diabetes, cancer, a cardiovascular disease, or a peripheral or central nervous system disorder.

18. A cDNA encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14.

19. The cDNA of claim 18 which comprises SEQ ID NO:1, 7 or 13.

20. The cDNA of claim 18 which consists of SEQ ID NO:1, 7 or 13.

21. An expression vector comprising a polynucleotide which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14.

22. The expression vector of claim 21 wherein the polynucleotide consists of SEQ ID NO:1, 7 or 13.

23. A host cell comprising an expression vector which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14.

24. The host cell of claim 23 wherein the polynucleotide consists of SEQ ID NO:1, 7 or 13.

25. A purified polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14.

26. The purified polypeptide of claim 25 which consists of the amino acid sequence shown in SEQ ID NO:2, 8 or 14.

27. A fusion protein comprising a polypeptide having the amino acid sequence shown in SEQ ID NO:2, 8 or 14.

28. A method of producing a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14, comprising the steps of:

culturing a host cell comprising an expression vector which encodes the polypeptide under conditions whereby the polypeptide is expressed; and isolating the polypeptide.

29. The method of claim 28 wherein the expression vector comprises SEQ ID NO:1, 7 or 13.

30. A method of detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14, comprising the steps of:

hybridizing a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO:1, 7 or 13 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and detecting the hybridization complex.

31. The method of claim 30 further comprising the step of amplifying the nucleic acid material before the step of hybridizing.

32. A kit for detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14, comprising:

a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO:1, 7 or 13; and instructions for the method of claim 30.

33. A method of detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14, comprising the steps of:

contacting a biological sample with a reagent that specifically binds to the polypeptide to form a reagent-polypeptide complex; and detecting the reagent-polypeptide complex.

34. The method of claim 33 wherein the reagent is an antibody.

35. A kit for detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14, comprising:

an antibody which specifically binds to the polypeptide; and instructions for the method of claim 33.

36. A method of screening for agents which can modulate the activity of a human patched-like protein, comprising the steps of:

contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO:2, 8 or 14 and (2) the amino acid sequence shown in SEQ ID NO:2, 8 or 14; and

detecting binding of the test compound to the polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential agent for regulating activity of the human patched-like protein.

37. The method of claim 36 wherein the step of contacting is in a cell.

38. The method of claim 36 wherein the cell is *in vitro*.

39. The method of claim 36 wherein the step of contacting is in a cell-free system.

40. The method of claim 36 wherein the polypeptide comprises a detectable label.

41. The method of claim 36 wherein the test compound comprises a detectable label.

42. The method of claim 36 wherein the test compound displaces a labeled ligand which is bound to the polypeptide.

43. The method of claim 36 wherein the polypeptide is bound to a solid support.

44. The method of claim 36 wherein the test compound is bound to a solid support.

45. A method of screening for agents which modulate an activity of a human patched-like protein, comprising the steps of:

contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 29% identical to the amino acid sequence shown in SEQ ID NO:2, 8 or 14 and (2) the amino acid sequence shown in SEQ ID NO:2, 8 or 14; and

detecting an activity of the polypeptide, wherein a test compound which increases the activity of the polypeptide is identified as a potential agent for increasing the activity of the human patched-like protein, and wherein a test compound which decreases the activity of the polypeptide is identified as a potential agent for decreasing the activity of the human patched-like protein.

46. The method of claim 45 wherein the step of contacting is in a cell.

47. The method of claim 45 wherein the cell is *in vitro*.

48. The method of claim 45 wherein the step of contacting is in a cell-free system.

49. A method of screening for agents which modulate an activity of a human patched-like protein, comprising the steps of:

contacting a test compound with a product encoded by a polynucleotide which comprises the nucleotide sequence shown in SEQ ID NO:1, 7 or 13; and

detecting binding of the test compound to the product, wherein a test compound which binds to the product is identified as a potential agent for regulating the activity of the human patched-like protein.

50. The method of claim 49 wherein the product is a polypeptide.

51. The method of claim 49 wherein the product is RNA.

52. A method of reducing activity of a human patched-like protein, comprising the step of:

contacting a cell with a reagent which specifically binds to a product encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO:1, 7 or 13, whereby the activity of a human patched-like protein is reduced.

53. The method of claim 52 wherein the product is a polypeptide.

54. The method of claim 53 wherein the reagent is an antibody.

55. The method of claim 52 wherein the product is RNA.

56. The method of claim 55 wherein the reagent is an antisense oligonucleotide.

57. The method of claim 56 wherein the reagent is a ribozyme.

58. The method of claim 52 wherein the cell is *in vitro*.

59. The method of claim 52 wherein the cell is *in vivo*.

60. A pharmaceutical composition, comprising:

a reagent which specifically binds to a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14; and a pharmaceutically acceptable carrier.

61. The pharmaceutical composition of claim 60 wherein the reagent is an antibody.

62. A pharmaceutical composition, comprising:

a reagent which specifically binds to a product of a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO:1, 7 or 13; and a pharmaceutically acceptable carrier.

63. The pharmaceutical composition of claim 62 wherein the reagent is a ribozyme.

64. The pharmaceutical composition of claim 62 wherein the reagent is an antisense oligonucleotide.

65. The pharmaceutical composition of claim 62 wherein the reagent is an antibody.

66. A pharmaceutical composition, comprising:

an expression vector encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2, 8 or 14; and a pharmaceutically acceptable carrier.

67. The pharmaceutical composition of claim 66 wherein the expression vector comprises SEQ ID NO:1, 7 or 13.

68. A method of treating a patched-like protein dysfunction related disease, wherein the disease is selected from diabetes, cancer, a cardiovascular disease, or a peripheral or central nervous system disorder comprising the step of:

administering to a patient in need thereof a therapeutically effective dose of a reagent that modulates a function of a human patched-like protein, whereby

symptoms of the patched-like protein dysfunction related disease are ameliorated.

69. The method of claim 68 wherein the reagent is identified by the method of claim 36.

70. The method of claim 68 wherein the reagent is identified by the method of claim 45.

71. The method of claim 68 wherein the reagent is identified by the method of claim 49.

Fig. 1

atgaaatcacaagccacacaatcactcagtgaaactagcgggactgctaaggt
cctttcccatttactgagagagccacaggaaccgagaggagaatatcagcatc
ctcttatcttgagccacactgcacagtcccttctcactccactgatgtcaggt
cctccagcagccttcagcactttgtgggattggcagagcagaagcagtcaccg
gtctttgctcacactcacagcttcagggcatgaacacacatcacactctctca
agaaacctctcaccatgcactggatgataacactgaaacaggattcagattct
gcagctcggaaatctttcagcatctttaccaattcagtcaatgtcatcactga
ccatggggttgccaaggctcaaaacctggacatcatcctttactcctttcctc
caacttcattgtactgtgttacacaaggtcctatatggaaaaagaatccacag
tggaattggaagtccaggctccctgctcaacctttgccagcagcacaggggaa
cagagggcattctgttcagaaaaaaatccagctgtggaatgtttttatttcat
cttcattgggaagaaaggccacaaaacattcagggatgccccttaaaacttgt
cctttgaagcattatgcttttatcaagcatctgtgctacagctttgaagattt
ttctcttgaatcatatttagtagaaatcaaagctgtgcataatctaagtttgc
aaagtcatggaactaaaggagtgtttgagcttctgtccggatggcggagaacc
aaagagaacttgcccttcaaagacaggatagcagatgcctattctgatgtgat
ggtcacctataccatgaccagctccctgtacttcatcacttttggcatgggtg
ccagcccattcacaaacatagaggctgtgaaggtcttctgtcaaaacatgtgt
gtctctattctgttgaactacttctacattttctccttctttggctcctgtct
ggtctttgctggccaactagagcaaaaccgctaccacagcatcttttgctgta
agatcccttctgcagaatacctggatcgcaaacctgtgtggttccagacagtg
atgagtgatgggcatcaacagacgtcccatcatgagacgaacccctaccagca
ccacttcattcagcacttcctccgtgaacattataatgaatggattaccaata
tatatgtgaagccatttgttgtcatcctctatctcatttatgcctccttctcc
ttcatggggtgcttacagatcagtgacggagccaacatcatcaatctactagc
cagtgattcgccaagtgtttcctatgccatggttcagcagaaatatttcagca
actatagccctgtgataggattctacgtctatgagcccctagagtactggaac
agcagcgtccaggatgacctaagaagactctgtagtggattcactgcagtgtc
ctgggtggagcagtactaccagttcctgaaagtcagcaacgtcagtgccaata
acaaaagtgacttcatcagtgtcctgcaaagctcatttttaaaaaagccagaa
ttccagcattttcgaaatgatatcatcttctccaaggcaggggatgaaagcaa
tatcattgcttctcgcttgtatctggtggccaggactagcagagacaagcaga
aagaaatcacagaagtgttggaaaagctgaggcccctatccctctcaaagagc
atccgattcatcgtgttcaacccctcctttgtcttcatggaccattacagctt
gtctgtcacagtgcctgttctgattgcaggctttggtgttctcctggtgttaa
tcctgacttttttcctagtgatccaccctctgggaaacttctggctaattctt
agcgtcacctcaattgagctgggcgttctgggcttaatgacattatggaacgt
cgacatggattgcatttctatcttgtgccttatctacaccttgaatttcgcca
ttgaccactgtgcaccactgcttttcacatttgtattagcaactgagcacacc
cgaacacaatgtataaaaagctccttgcaagaccatgggacagccattttgca
aaatgttacttctttttcttattgggttagtcccccttctatttgtgccttcga
acctgaccttcacactgttcaaatgcttgctgctcactgggggttgcacactt
ctgcactgttttgttatttttacctgtgttcctaacgtttttcccccccttccaa
aaagcaccacaagaaaaagaaacgtgccaagcgaaaggagagagaggaaattg
aatgcatagaaattcaagagaacccggatcacgtcaccacagtatga

Fig. 2

MKSQATQSLSETSGTAKVLSHLLREPQEPRGEYQHPLILSHTAQSLLTPLMSG
PPAAFSTLWDWQSRSSHRSLLTLTASGHEHTSHSLKKPLTMHWMITLKQDSDS
AARKSFSIFTNSVNVITDHGVAKAQNLDIILYSFPPTSLYCVTQGPIWKKNPQ
WNWKSRLPAQPLPAAQGNRGHSVQKKIQLWNVFISSSLGRKATKHSGMPLKTC
PLKHYAFIKHLCYSFEDFSLESYLVEIKAVHNLSLQSHGTKGVFELLSGWRRT
KENLPFKDRIADAYSDVMVTYTMTSSLYFITFGMGASPFTNIEAVKVFCQNMC
VSILLNYFYIFSFFGSCLVFAGQLEQNRYHSIFCCKIPSAEYLDRKPVWFQTV
MSDGHQQTSHHETNPYQHHFIQHFLREHYNEWITNIYVKPFVVILYLIYASFS
FMGCLQISDGANIINLLASDSPSVSYAMVQQKYFSNYSPVIGFYVYEPLEYWN
SSVQDDLRRLCSGFTAVSWVEQYYQFLKVSNVSANNKSDFISVLQSSFLKKPE
FQHFRNDIIFSKAGDESNIIASRLYLVARTSRDKQKEITEVLEKLRPLSLSKS
IRFIVFNPSFVFMDHYSLSVTVPVLIAGFGVLLVLILTFFLVIHPLGNFWLIL
SVTSIELGVLGLMTLWNVDMDCISILCLIYTLNFAIDHCAPLLFTFVLATEHT
RTQCIKSSLQDHGTAILQNVTSFLIGLVPLLFVPSNLTFTLFKCLLLTGGCTL
LHCFVILPVFLTFFPPSKKHHKKKKRAKRKEREEIECIEIQENPDHVTTV

Fig. 3

MIRREVFQELRQLDNIIQNATTTYDGDTYTYKDNCARWENECFENDILNLDAL
MDDIEAGQLNLTFPFMFNPVTWDAHLFPVFFGGTKLTEDNYVISVPAIQLVYF
VTADTKRQDAKGAEWEETFLRVVGNAENSGQFKHISVSYFASRTLDHELEKNT
KTVVPYFSSTFLLMGLFSIITCMMGDAVRSKPFLGLMGNVSAIMATLAAFGLA
MYCGIEFIGINLAAPFLMIGIGIDDTFVMLAGWRRTKAKMPVAERMGLMMSEA
AVSITITSVTDFISFLIGIISPFRSVRIFCTYSVFAVCFTFLWHITFFAACMA
ISGYRERKNLHSIFGCRVQPMSVAIKEKRNFLYKAIMAGGIDANDPDNPIDNK
DHMLMAFFKDKMAAVINNKWCKAIIILAFASYLVGACYGITQIKEGLERRKLS
REDSYSVEFFDREDDYYREFPYRMQVIIAGPLNYSDPLVQEQVENLTSTLEHT
SYVTSRRYTESWLRSFLSFLERNNELLNVTVDDEQTFIDAVKEHWLFPGNPFS
LDVRFNEDETQIIASRFLIQAVNITDTNHEKEMVRDLRQICKDSPLNASIFHP
YFVFFDQFELVRPVSLQAMVIGAIIMMIISFVFIPNILCSLWVAFSVISIELG
VAGYMALWDVNLDSISMINLIMCIGFSVDFTAHICYTYMSSKKRSPKARVREA
LHSLGLPIIQGSSSTILGIVALLLAQSYIFLVFFKMVFLVIFFGAMHGLFLLP
VLLSLFGPGSCLTWTGKDDGSDAEVDDGLDDRQLEKPFSQSYYMQYPSIGING
PYGSKGFLGAPYKAYGVDEKDLGLGTSGEDSSESSSSRSQHRQQAAATEEEVV
VRESPTRRYDDGWRRSSYQNIYGQGAAQFQAQPDLYGKQVSATEWRQRLDTHE
QQQRQRQRRSPFENYRQDVEIDMQKARRNSHGDVIDLHGTPNSSVEERFRRRG
EPFSAESGDDSSYRHQQIMAMPAAGSAPSAKRYHRRRSSEDSTSRHQRWPANI
EERRARRAYSPAHNRPETALTSYAYRSSSHHNLYQPNGKSSKYPPTYQYGDYY
H

Fig. 4

AGAATTATTGCACAGGGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACT
TTGGGAGGCCGAGGCGGGTGGATCATGAGGTCAGGAGATCGAGACCATCCTGG
CTAACAAGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGGCGC
GGTGGCGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGG
CGTGAACCCGGGAAGCGGAGCTTGCAGTGAGCCGAGATTGCGCCACTGCAGTC
CGCAGTCCGACCTGGGCGACAGAGCGAGACTCCGTCTCAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

Fig. 5

CCATGGGAAACAAACAGAGAAGAACTCTATGATCTCTTGGAAACCCTGAGGAG
ACTTTCTGTCACCTCCAAGGTGAAGTTCATCGTCTTCAATCCGTCCTTGTATA
CATGGATCGATATGCCTCCTCTCTGGGAGCCCCCCTGCACAACTCCTGCATCA
GTGCTTGTTCCTGCTCTTCTTCTCGGCATTCCTGGTGGCAGATTCACTGATTA
ACGTCTGGATCACTCTCACAGTTGTGTCCGTGGAGTTGGAGTGATAGGTTTCA
TGACATTATGGAAAGTAGAACTGGACTGCATTTCTGTGCTATGCTTAATTTAT
GGAATTAATTACACAATTGACAATTGTGCTCCAATGTTATCCACATTTGTTCT
GGGCAAGGATTTCACAAGAACTAAATGGGTAAAAAATGCCCTGGAAGTGCATG
GGGTAGCTATTTTACAGAGTTACCTCTGCTATATGTTGGTCTGATTCCTCTTG
CAGCTGTGCCTTCAAATCTGACCTGTACACTGTTCAGGTGCTTGTTTTTAATA
GCATTGTCACCTTCTTTCACTGCTTGCCATTTTACCTGTGATACTGACTTTCC
TGCCACCCTCTAAGAAAAAAAGGAAAGAGAAGAAAAATCCTGAGAACCGGGAG
GAAATTGAGTGTGTAGAAATGGTAGATATCGATAGTACCCGTGTGGTTGACCA
AATTACAACAGGTGATATGTCTGCTTGGATATTTCACCTTAGGTTTATCAGAA
CAAGAGATTTGTTATGAAACATTAATTCCAAGGTCTTCCCTTTAAAGATTGAA
CGGTTGGCAAAAAAAAAAAAATGGGGTTTTTGGGGGGAAAGGTTAAGGGGGAA
GGTCGCCTTGAAGGGAAGGGCCTAG

Fig. 6

GACTGATCGACTTCAGACATCATGGAAACTCGCACTGGATTCCATTTCTTTGC
TATGCTTTATTTATGGGATTCACTACACAATTGACCACTGTGCTCCACTTTCA
TCCACATTTGTTCTAGGCAAGGATTTTACAAGAACTAAATGGGTTTAAAATCC
CCTGGAAGTGCATGGGGTAGCTATTTTACAGAGTTACCTCTGTTATATTGTTG
GTCTATTTCCTCTTGCAGCTGTCCCTTCAAATCTGACCTGTACACTGTTCAGG
TGCTTGTTTTTAATAGCATTTGTCACCTTCTTTCACTGCTTTGCCATTTTACC
TGTGATACTTACTTTCCTGCCACCCTCC

Fig. 7

BLASTP - alignment of SEQ ID NO:7 against trembl|AE003784|AE003784_10

gene: "CG11212"; Drosophila melanogaster genomic scaffold 142000013386052 section 1 of 5, complete sequence.//:gp|AE003784|7302178 gene: "CG11212"; Drosophila melanogaster genomic scaffold 142000013386052 section 1 of 5, complete
sequence.

Pfam HMM search predict this Drosophila protein belong to the Patched family from
residues 1 to 957 with score 556.2 and E-value 9e-165.

This hit is scoring at : 6e-56 (expectation value)
Alignment length (overlap) : 525
Identities : 26 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb

```
Q:    251 GTKGVFELLSGWRRTKENLPFKDRIADAYSDVMVTYTMTSSLYFITFGMGA-SPFTNIEA
          G....F :L:GWRRTK..:P..:R:. ..S:..V:.T:TS . FI:F :G. SPF   .:
H:    234 GIDDTFVMLAGWRRTKAKMPVAERMGLMMSEAAVSITITSVTDFISFLIGIISPF---RS

          VKVFCQNMCVSILLNYFYIFSFFGSCLVFAGQLEQNRYHSIFCCKIP--SAEYLDRKPVW
          V::FC. ...:....:.: .:FF.:C:..:G. E:...HSIF C::.  S... ::: .
          VRIFCTYSVFAVCFTFLWHITFFAACMAISGYRERKNLHSIFGCRVQPMSVAIKEKRNFL

          FQTVMSDGXQQTSXXETNPYQXXFIQXFLREXYNEWITNIYVKPFVVILYLIYASFSFMG
          ::.:M:.G..... ... . :...:..F.::.. . I.N :.K..:::.:. Y.  : .G
          YKAIMAGGIDANDPDNPIDNKDHMLMAFFKDKMAAVINNKWCKAIIILAFASYLVGACYG
```

Fig. 7 (continued)

```
CLQISDGANIINLLASDSPSVSYAMVQQKYFSNYSPVIGFYVYEPLEYWNSSVQDDLRRL
..QI.:G..  .L ..DS SV.:   :..Y:..:.  : ..:  PL.Y :. VQ:.:..L
ITQIKEGLERRKLSREDSYSVEFFDREDDYYREFPYRMQVIIAGPLNYSDPLVQEQVENL

CSGFTAVS----------WVEQYYQFLK----VSNVSANNKSDFISVLQSSFLKKPEFQH
.S .. .S          W:..:..FL:    : NV:.:::..FI..:..:L . .
TSTLEHTSYVTSRRYTESWLRSFLSFLERNNELLNVTVDDEQTFIDAVKEHWLFPGN--P

FRNDIIFSKAGDESNIIASRLYLVARTSRDKQKEITEVLEKLRPLSLSKSIRFIVFNPSF
F. D: F::  DE:.IIASR..: A..D...E .E:..LR.:. ...:.  :F:P F
FSLDVRFNE--DETQIIASRFLIQAVNITDTNHE-KEMVRDLRQICKDSPLNASIFHPYF

VFMDHYSLSVTVPVLIAGFGVLLVLILTFFLVIHPLGNFWLILSVTSIELGVLGLMTLWN
VF.D.:.L  .V.: .. .G.::::I::F..: : L :.W:..SV.SIELGV.G.M.LW:
VFFDQFELVRPVSLQAMVIGAIIMMIISFVFIPNILCSLWVAFSVISIELGVAGYMALWD

VDMDCISILCLIYTLNFAIDHCAPLLFTFVLATEHTRTQCIKSSLQDHGTAILQNVTSFL
V::D.IS:: LI..:.F::D..A :.:T:: :.:.: .. ::.:L.. G..I:Q. :S :
VNLDSISMINLIMCIGFSVDFTAHICYTYMSSKKRSPKARVREALHSLGLPIIQGSSSTI

IGLVPLLFVPSNLTFTLFKCLLLTGGCTLLHCFVILPVFLTFFPP     758
:G:V.LL...S : ...FK.:.L.    .:H ..:LPV.L:.F P
LGIVALLLAQSYIFLVFFKMVFLVIFFGAMHGLFLLPVLLSLFGP     750
```

Fig. 8

HMMPFAM - alignment of SEQ ID NO:7 against pfam|hmm|Patched

Patched family -

This hit is scoring at : -179.3
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:      38 ILSHTAQSLLTPL-MS---GPPAAFSTLWDWQSR-------SSHRSLLTLTASGHEH---
           :LS...  ..TP :.    . .. S.  .  .R       .. ...:.:TA.....
H:       1 vlsskilytftpsDirksytergvrsedeplverkFfpgdgkpiarfviitakarnsedy

          TSHSLKKPLTMHWMITLKQDSDSAARKSFSI----FTNSVNVITDHG------VAKAQN-
          T. :: .   :: ::.:..  : .. ..S     .T S.. :...        V : .N
          trpNiLdikyLnevvqvndfiltnvtipvsgalkkitlsfkdlCssfcdincpvnqhyng

          LDI--ILYSFPPTSLYCVTQGPIWKKNPQWNWKSRLPAQPLPA----------------
          LD.  IL S   .S . :T. .:   ..:    ... . .L..
          Ldtlailrsnlhnsridl tYPtmtffgtkiylgpnfgGVkldpgeNdvekllDsltklkl

          -----AQGNRGHSvqKKIQLWnvFISSSLGRKATKHSG---MPLKTCPLKHYAFIKHL-C
               .:GN  .   K.I L:  :  ..  .:..:.S    :.L:.   ..Y.  .H: .
          vdtkseegniksv..kaivLy..yrlkrdpeeveedskeWElsledyleneyksdshiev

          YSFEDFSLESYLVE---------IKAVHNLSLQSHGT-----K----------------
            F.D  LE. LV.          : ..  L . S  T
          tifsdqvledElvrngltlvPffvvgFaiLvtFsvltsvvlavrrsgslhidwvssKpiL

          ------------------------------------G------VFELLSGWRRTKENL
                                              G      :F ::..W:RT..:L
          AilgvltPlmAivsAfGlLfwlGfrFnsIlcVmPFLVLGAIGVDDmFLMvhaWqrttrsl
```

Fig. 8 (continued)

```
P------FKDRIADAYSDVMVTYTMTSSLYFITFGMGAspFTNIEAVKVFCQNMCVSILL
.      .K.R:.:.. :.  :.T:TS . .::FG:GA  .T ...:::FC  ..V:I..
ssrarKvvkkRmgevlvEaGPSItITSlTnvLSFgIGa..iTptPeIqlFCiytavAiff

NYFYIFSFFGSCLVFAGQLEQNRYHSIFCC----KIPSAEylDRKPVWF--QTVMSDGHQ
::.Y .:FF.:.L..AG:.E..  HS:F.     .:.:AE  ...P  .  Q.:.S G..
dfiYQlTFFaAilaiagkyEmkgrhslflriTakavdaae..tespkklkkqrlksrgsa

Q-------------------------TSHHETNPYQHH--FIQHFLREHYNEWITNIYV
:                          .:S....:  ...   : .F. . Y .::.N  V
kaaPvekegeknSkfseveeslesekssdssksheaeqktavvkfflniYcpfllnpkv

KPFVVILYLIYASFSFMGCLQISDGANIINLLASDSPSVSYAMVQ-QKYFSNYSPVIgFY
:  V:::Y::Y.:.:..GC..:.:G.:  .LL..DSP V.....:: :..:.  S.V. ..
rvcvllvyvvYlaiaiYGctnlkegLdPskLllkDSpLveilrlrekhvwpyGsqvt.Vv

VYEPLEYWNSSVQDDLRRLCSGF--------T----AVSWVEQYYQFL--------KVSN
V .P :. N.. :D L .:.. F         .    .  W:..Y :FL        :.S.
VNNPPDLtnpenrdrllemvdeFEntpyalGkWEnsTkfWLrdYekFlsssEeinlease

VSANNK----------------SDFISVLQSSFLKKPEFQHFRNDIIFS-KAGDESNIIA
:. ..:                . ..  L . FLK.P F.H::.D:::. K...E:. I.
ledeeeeidevdpvsvvvkdkqewifydllewFlkspgfshWqgdlvwdnktdketteik

SRLYLVA---------RTSRDKQKEITEVLEKLRP-LSLSksirfiVFNPSFVFMDHYSL
.  :...          .....  KE . V.::. . .:::      VF:.. .F:D..
kFrFttgyHgkdlstwtdRtkllkewRgvAdeypddFNVT......VFdedafFlDqils

SVTVPVLIAGFGVLLVLILTFFLVIHPLGNFWLILSVTSIELGVLGLMTLWNVDMDCISI
 :...:  . : :::.:::.F..: :P:  F :.:S:.SI::GV.G.::LW.VD:D IS:
ilpttiQsiiwtlicMavVCflFIpnpitvfvitvsiaSIdiGVfGfLSLWGVdLDPisM
```

Fig. 8 (continued)

```
LCLIYTLNFAIDHCAPLLFTFVLATEH----------TRTQCIKSSLQDHGTAILQNVT
:.:I.:::.F::D..A :.: F. :. .           T .: :..:L:  G .::Q ..
itiIMSIGFSVDfsAHiaYhFyrskgsesrsikkiplktpdeRvadALealGWPviQAal

SFLIGLVPLLFVPSNLTFTLFKCLLLTGGCTLLHCFVILPVFLTFFPP----------S
S :: :: LLFVPS :....FK.:.L.  . LLH .:.LP:.L:.F .          .
STiLcvlvLlFVpsYmvvvFfKTifLVVvlGlLHGLifLPilLslfvtssrifggdviir

KKHHKKKKRAKR    771
..: K KK :::
tsnakikkpseq    957
```

Fig. 9

atgctcaatgaagaaaaagtggatttttccattgctggtcatggattatatgg
gacttttgaaatgttatcctcctggaggaaaactagagaagaccaacatgtta
aagagagaactgcagcagtctatgcagactccatgctctccttttctctcacc
actgccatgtacctggtcacctttggcataggggccagcccttcacgaacat
tgaggcagccaggattttctgctgcaattcctgtattgcaatcttcttcaact
acctctatgtactctcgttttatggttccagcctagtgttcactggctacata
gaaaacaattaccagcatagtatcttctgtagaaaagtcccaaagcctgaggc
attgcaggagaagccggcatggtacaggtttctcctgacggccagattcagtg
aggacacagctgaaggcgaggaagcgaacacttacgagagtcacctattggta
tgtttcctcaaacgctattactgtgactggataaccaacacctatgtcaagcc
ttttgtagttctcttttaccttatttatatttcctttgccttaatgggctatc
tgcaggtcagtgaagggtcagaccttagtaacattgtagcaaccgcgacacaa
accattgagtacactactgcccagcaaaagtacttcagcaactacagtcctgt
gattgggttttacatatatgagtctatagaatactggaacactagtgtccaag
aagatgttctagaatacaccaaggggtttgtgcggatatcctggtttgagagc
tatttaaattaccttcggaaactcaatgtatccactggcttgcctaagaaaaa
tttcacagacatgttgaggaattccttttctgaaagcccctcaattttcacatt
ttcaagaggacatcatcttctctaaaaaatacaatgatgaggtcgatgtagtg
gcctccagaatgtttttggtggccaagaccatggaaacaaacagagaagaact
ctatgatctcttggaaaccctgaggagactttctgtcacctccaaggtgaagt
tcatcgtcttcaatccgtcctttgtatacatggatcgatatgcctcctctctg
ggagccccccctgcacaactcctgcatcagtgctttgttcctgctcttcttctc
ggcattcctggtggcagattcactgattaacgtcggatcactctcacagttgt
gtccgtggagtttggagtgataggtttcatgacattatggaaagtagaactgg
actgcatttctgtgctatgcttaatttatggaattaattacacaattgacaat
tgtgctccaatgttatccacatttgttctgggcaaggatttcacaagaactaa
atgggtaaaaaatgccctggaagtgcatggggtagctattttacagagttacc
tctgctatattgttggtctgattcctcttgcagctgtgccttcaaatctgacc
tgtacactgttcaggtgcttgtttttaatagcatttgtcaccttctttcactg
ctttgccattttacctgtgatactgactttcctgccaccctctaagaaaaaaa
ggaaagagaagaaaaatcctgagaaccgggaggaaattgagtgtgtagaaatg
gtagatatcgatagtacccgtgtggttgaccaaattacaacagtgtga

Fig. 10

MLNEEKVDFSIAGHGLYGTFEMLSSWRKTREDQHVKERTAAVYADSMLSFSLT
TAMYLVTFGIGASPFTNIEAARIFCCNSCIAIFFNYLYVLSFYGSSLVFTGYI
ENNYQHSIFCRKVPKPEALQEKPAWYRFLLTARFSEDTAEGEEANTYESHLLV
CFLKRYYCDWITNTYVKPFVVLFYLIYISFALMGYLQVSEGSDLSNIVATATQ
TIEYTTAQQKYFSNYSPVIGFYIYESIEYWNTSVQEDVLEYTKGFVRISWFES
YLNYLRKLNVSTGLPKKNFTDMLRNSFLKAPQFSHFQEDIIFSKKYNDEVDVV
ASRMFLVAKTMETNREELYDLLETLRRLSVTSKVKFIVFNPSFVYMDRYASSL
GAPLHNSCISALFLLFFSAFLVADSLINVWITLTVVSVEFGVIGFMTLWKVEL
DCISVLCLIYGINYTIDNCAPMLSTFVLGKDFTRTKWVKNALEVHGVAILQSY
LCYIVGLIPLAAVPSNLTCTLFRCLFLIAFVTFFHCFAILPVILTFLPPSKKK
RKEKKNPENREEIECVEMVDIDSTRVVDQITTV

Fig. 11

CTCCTATTCTTGATTAATGGAATCGACTACACAACTGACAACTCTGCTCCACT
GTTATCCACATTTCTTCTAGGCAAGGATTTCACCAGAACTAAATGGGTTAAAA
ATCCCCTGGAAGTGCATGGGGTAGCTATTTTACAGAGTTACCTCTGTTATATT
GTTGGTCTATTTCCTCTTGCAGCTGTCCCTTCAAATCTGACCTGTACACTGTT
CAGGTGCTTGTTTTTAATAGCATTTGTCACCTTCTTTCACTGCTTTGCCATTT
TACCTGTGATACTTACTTTCCTCCCACCCTC

Fig. 12

GCTTAATTTATGGAATTAACTACACAATTGACAACTCTGCTCCATTGTCTTCC
AAATTTGTTCCAGGCAAGGATTTTACAAGAACTAAATGGGTTAAAAATGCCCT
GGAAGTGCATGGGGTAGCTATTTTACAGAGTTACCTCTGTTATATTGTTGGTC
TATTTCTTCTTGCAGCTGTGCCTTCAAATCTGACCTGTACACTGTTCAGGTGC
TTGTTTTTAATAGCATTTGTCACCTTCTTTCACTGCTTTGCCATTTTACCTGT
GATACTTACTTTCCTGCCACCCTCG

Fig. 13

BLASTP - alignment of SEQ ID NO:13 against trembl|AE003784|AE003784_10

gene: "CG11212"; Drosophila melanogaster genomic scaffold 142000013386052 section 1 of 5, complete sequence.
//:gp|AE003784|7302178 gene: "CG11212"; Drosophila melanogaster genomic scaffold 142000013386052 section 1 of 5, complete sequence.

Pfam HMM search predict this Drosophila protein belong to the Patched family from
residues 1 to 957 with score 556.2 and E-value 9e-165.

This hit is scoring at : 5e-56 (expectation value)
Alignment length (overlap) : 536
Identities : 27 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb

```
Q:       9 FSIAGHGLYGTFEMLSSWRKTREDQHVKERTAAVYADSMLSFSLTTAMYLVTFGIGA-SP
            F :.G G: .TF ML:.WR:T:... V.ER...:.:::.:S.::T:.. .::F IG. SP
H:     228 FLMIGIGIDDTFVMLAGWRRTKAKMPVAERMGLMMSEAAVSITITSVTDFISFLIGIISP

            FTNIEAARIFCCNSCIAIFFNYLYVLSFYGSSLVFTGYIENNYQHSIF-CRKVPKPEALQ
            F.::   RIFC. S..A: F.:L: ::F:.:.:..:GY E..  HSIF CR  P.. A::
            FRSV---RIFCTYSVFAVCFTFLWHITFFAACMAISGYRERKNLHSIFGCRVQPMSVAIK

            EKPAW-YRFLLTARFSEDTAEGEEANTYESHLLVCFLKRYYCDWITNTYVKPFVVLFYLI
            EK  : Y: ::.. ...:.:.:..  N  :.H:L:.F.K  ..  I.N.:.K..::L :.
            EKRNFLYKAIMAGGIDANDPDNPIDN--KDHMLMAFFKDKMAAVINNKWCKAIIILAFAS
```

Fig. 13 (continued)

```
YISFALMGYLQVSEGSDLSNIVATATQTIEYTTAQQKYFSNYSPVIGFYIYESIEYWNTS
 Y:  A..G..Q:.EG : ..: .. :.::E: ..:..Y:..:.  : ..I  .:.Y :.
 YLVGACYGITQIKEGLERRKLSREDSYSVEFFDREDDYYREFPYRMQVIIAGPLNYSDPL

 VQEDV------LEYTKGFVRI----SWFESYLNYLRK----LNVSTGLPKKNFTDMLRNS
 VQE.V      LE:T. ...      SW..S:L::L.:    LNV:.. .::.F.D.::..
 VQEQVENLTSTLEHTSYVTSRRYTESWLRSFLSFLERNNELLNVTVD-DEQTFIDAVKEH

 FLKAPQFSHFQEDIIFSKKYNDEVDVVASRMFLVAKTM-ETNREELYDLLETLRRLSVTS
 :L .   : F. D: F::   DE..::ASR..: A .: :TN.E:  ::::.LR::. .S
 WLFPG--NPFSLDVRFNE---DETQIIASRFLIQAVNITDTNHEK--EMVRDLRQICKDS

 KVKFIVFNPSFVYMDRYASSLGAPLHNSCISALFLLFFSAFLVADSLINVWITLTVVSVE
 .:.  :F:P FV:.D::.    ..L. ..I.A:.::.:S ..:.: L.::W:..:V:S:E
 PLNASIFHPYFVFFDQFELVRPVSLQAMVIGAIIMMIISFVFIPNILCSLWVAFSVISIE

 FGVIGFMTLWKVELDCISVLCLIYGINYTIDNCAPMLSTFVLGKDFTRTKWVKNALEVHG
 .GV.G:M.LW.V.LD.IS:: LI. I.:::D .A :. T:: .K. : .. V:.AL.  G
 LGVAGYMALWDVNLDSISMINLIMCIGFSVDFTAHICYTYMSSKKRSPKARVREALHSLG

 VAILQSYLCYIVGLIPLAAVPSNLTCTLFRCLFLIAFVTFFHCFAILPVILTFLPP     526
 :.I:Q.  . I:G::.L....S : ...F:.:FL:.F.  .H . :LPV:L:.. P
 LPIIQGSSSTILGIVALLLAQSYIFLVFFKMVFLVIFFGAMHGLFLLPVLLSLFGP     750
```

Fig. 14

HMMPFAM - alignment of SEQ ID NO:13 against pfam|hmm|Patched

Patched family -

This hit is scoring at : -357.9 and E-value 1.2e-07
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:      1 ----------------------------MLNEE-------------------------
                                      L E.
H:      1 vlsskilytftpsDirksytergvrsedeplverkFfpgdgkpiarfviitakarnsedy

          -------------------------KVDFSI-------------------------
                                  .:..S
          trpNiLdikyLnevvqvndfiltnvtipvsgalkkitlsfkdlCssfcdincpvnqhyng

          ------------------------------------AGHGLYG---------------
                                             .G  L
          LdtlailrsnlhnsridltyptmtffgtkiylgpnfgGVkldpgeNdvekllDsltklkl

          ---------------------------------------------TFE---------
                                                      . :
          vdtkseegniksvkaivLyyrlkrdpeeveedskeWElsledyleneyksdshievtifs

          ------------------------------------------------------------

          dqvledElvrngltlvPffvvgFaiLvtFsvltsvvlavrrsgslhidwvssKpiLAilg

          ------------------------------------------MLSSWRKTREDQH---
                                                   M:.:W::T... .
          vltPlmAivsAfGlLfwlGfrFnsIlcVmPFLVLGAIGVDDmFLMvhaWqrttrslssra
```

Fig. 14 (continued)

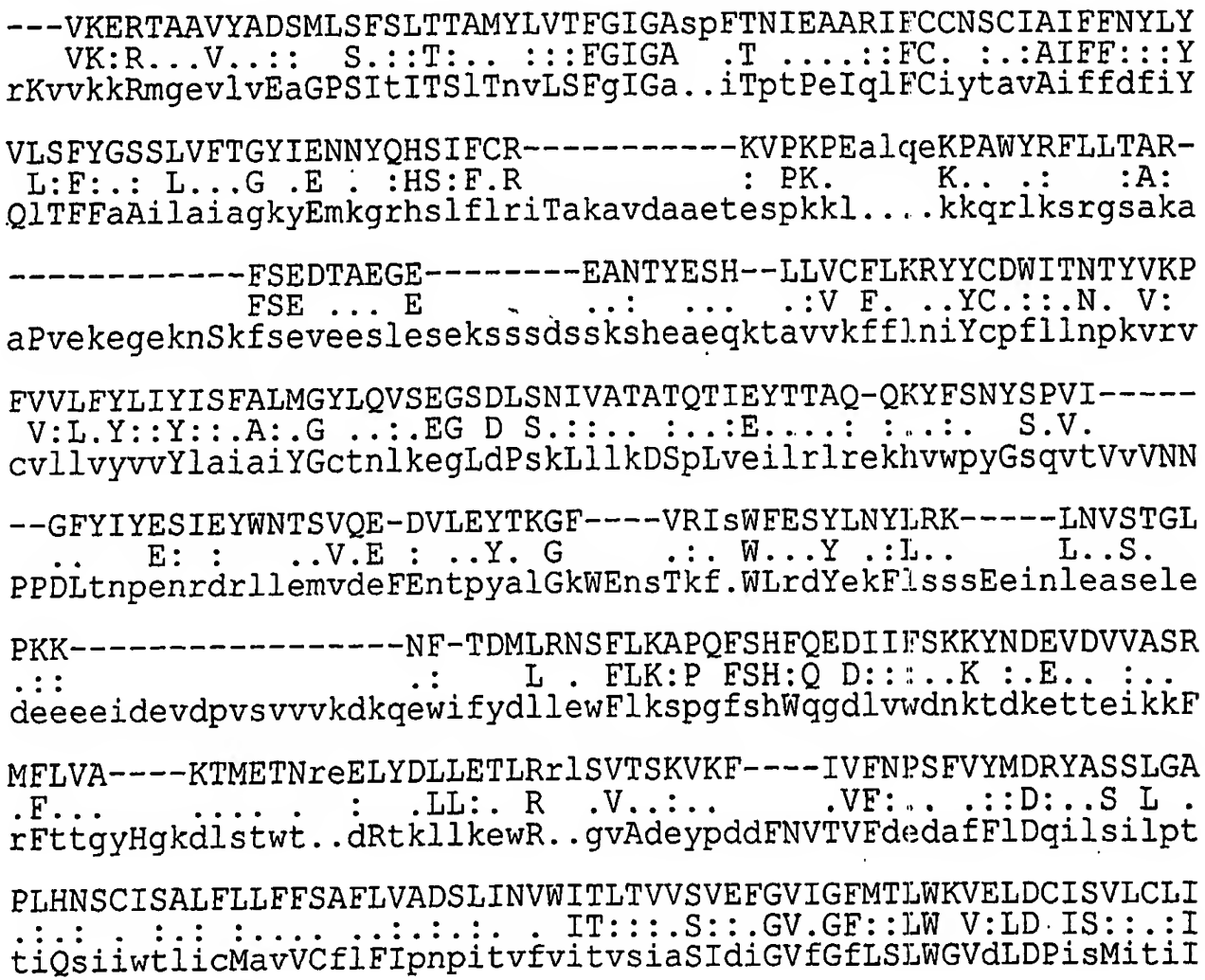

Fig. 14 (continued)

```
YGINYTIDNCAPMLSTFVLGKDF-----------TRTKWVKNALEVHGVAILQSYLCYIV
..I.:::D .A :.  F. .K.             T .: V.:ALE. G .::Q: L. I:
MSIGFSVDfsAHiaYhFyrskgsesrsikkiplktpdeRvadALealGWPviQAalSTiL

GLIPLAAVPSNLTCTLFRCLFLIAFVTFFHCFAILPVILTFLPP--------------SK
 :: L. VPS :....F:.:FL:..: ..H ...LP::L:.. .                :.
cvlvLlFVpsYmvvvFfKTifLVVvlGlLHGLifLPilLslfvtssrifggdviirtsna

KKRKEKKN     536
K :K..:.
kikkpseq     957
```

Fig. 15

atgccgtgggtggagcccaagcccaggccggggccggagcagaagccca
agctcaccaaaccggactctgccaccgggccgcagtggtaccaggaatc
tcaggaatcggagtcggaaggcaagcagccacccccgggacccctggca
cccccgaaatcccccgaaccctcaggacccctggcgtcggagcaggatg
cacccctgccagaggggacgatgcacccccccggccgtcgatgctgga
cgatgcaccccgcctgccgctggagctggacgatgcaccctgccggag
gaggaaaccccccgaacccacggccatctgcaggcaccggcaccgctgtc
acaccgactgcctagaggggctgctgtcccgcaccttccagtggctggg
gtggcaggtgggcgcgcacccctggatcttcctgctggcgcccttgatg
ctgacagccgcgctgggcaccggcttcctgtacctacccaaggacgaag
aggaagacctagaggagcattacacccctgtggggagcccggccaaggc
ggagcggcgcttcgtgcagggccatttcaccaccaacgactcctaccgc
ttctccgcctccaggaggagcaccgaagccaatttcgtctcgcttctgg
tggtctcctacagcgactcactgctggacccagctacctttgcagaagt
cagcaaactggacggcgcggtgcaggatctgcgcgtggcgcgggaaaag
ggaagccagatccagtaccagcaggtgtgcgcgaggtacagggcgctct
gcgtgccccccaacccgatcctgtacgcctggcaggtgaacaaaacgct
caacctgagcagcatctccttccccgcctacaaccacggcaggcatccc
ctctacctgaccggcttcttcggaggatacatcttgggggcagcctag
gaatgggccagttactcctgcgggccaaagccatgcggctgctgtacta
cctgaagaccgaggaccctgagtacgacgtgcagagcaagcagtggctc
acccatttgctcgatcaatttaccaacattaagaacatcttggccttga
aaaaaattgaggtacctggtggtgtgggtttacagggaggccaggagaa
ggtagtccactttacatcgctttccagacaactggaatttgaggcaact
tctgtgactgtgatccctgtgtttcacctggcatacattctcatcattc
tgtttgcagtcacatcatgctttagcttctctatttctgggggggaaat
ggttgttggaattttgatggggattgtgttgaatctgcagattgttttg
agtaggactggcatttttaaaaatattgagtctgccaacccaggaactaa
ggatctttccatttatccaggtcttttttaagtttcttttaccaatgttt
gctggttttcgatgaacatggaactgatatccatccaataagtttgttt
tttagagactattttggccccttttctcacaaggagtgagtccaagtatt
ttgtagtctttatatatgttttgtacatcataagcagtatatatgggtg
tttccatgtgcaggaaggtttagaccttcgaaatctggcaagtgacgat
tcctacatcacaccatattttaacgtagaggagaattatttttcagatt
atggtcccagggttatggttattgttactaaaaaagttgactactggga
taaagatgttaggcaaaaactggaaaactgtactaaaatttttgaaaaa
aatgtctatgtagataaaaatcttacagagttttggttagatgcatatg
tgcaatatttaaaaggattcatgaacaatattgtatgggagaaactgag
ctcatgcaactatgctatcaatcagacttggctggtgaaagccaatgca
tctatccccttgtatgggccactgaacaataaaatgaggaaaggtccag
gaggaattgtacatactaggatcttggtggaaaggctcacctgcctgct
gacatcagttctggcagtgaacctgaaagttgctctgtggctcagcatt
agccccctcagctgaggcccagctcagaactgctcacacaaggaccca
gagggacacttgcccatatctcacagctcaagagcctgagcttccctga
aagttttgccaacttctgtctcacagcagattccaaggaggcccggtct
cagctcctgctgccatcaggtaactgtcccatctatgctgaaacgtgtt
gggaaacaaagtgccccctgattcttcgatatgggctctccagcctcc
atcccacagcagatcccaaggggtccaattccagctctggcctctctt
gtt

Fig. 16

MPWVEPKPRPGPEQKPKLTKPDSATGPQWYQESQESESEGKQPPPGPLA
PPKSPEPSGPL ASEQDAPLPEGDDAPPRPSMLDDAPRLPLELDDAPLP
EEETPEPTAICRHRHRCHTDCLEGLLSRTFQWLGWQVGAHPWIFLLAPL
MLTAALGTGFLYLPKDEEEDLEEHYTPVGSPAKAERRFVQGHFTTNDSY
RFSASRRSTEANFVSLLVVSYSDSLLDPATFAEVSKLDGAVQDLRVARE
KGSQIQYQQVCARYRALCVPPNPILYAWQVNKTLNLSSISFPAYNHGRH
PLYLTGFFGGYILGGSLGMGQLLLRAKAMRLLYYLKTEDPEYDVQSKQW
LTHLLDQFTNIKNILALKKIEVPGGVGLQGGQEKVVHFTSLSRQLEFEA
TSVTVIPVFHLAYILIILFAVTSCFRFDCIRNKMCVAAFGVISAFLAVV
SGFGLLLHIGVPFVIIVANSPFLILGVGVDDMFIMISAWHKTNLAGDIR
ERMSNVYSKAAVSITITTITNILALYTGIMSSFSIYGCFHVQEGLDLRN
LASDDSYITPYFNVEENYFSDYGPRVMVIVTKKVDYWDKDVRQKLENCT
KIFEKNVYVDKNLTEFWLDAYVQYLKGFMNNIVWEKLSSCNYAINQTWL
VKANASIPLYGPLNNKMRKGPGGIVHTRILVERLTCLLTSVLAVNLKVA
LWLSISPPQLRPSSELLTQGPRGTLAHISQLKSLSFPESFANFCLTADS
KEARSQLLLPSGNCPIYAETCWETKCPSDSSIWALQPPSHSRSQGGPIP
AL

Fig. 17

MAWDCVERRAASLFRQLGFLICDHPLPFFVFPLLFTAAMGVGLLHLNPL
SDAVYLFTPLGAQSKMERMSIHEKWPLTDNNYIPGRAVTQSREIQVTAL
ARNDSNILDPKFANAVYQLDKYIQTRVRVLHNGHYYSYKNLCLQYKNGG
CPSNKHVHILSDLHNHGFNITYPYFRFGSEGGYIGSSLGGVTVMKGENE
TDILASAKAWFMIYHLKFHPEEMSYISGEWELELGRMLTQYPEDPYISI
TYFHSQTLADELKRNADTLIPRFIISITLLIVFSTLCSLSFIDGSFSID
WVLSKPILSILGVVSAGIAILTGVGFLSLMGMPYNDIVGVMPFLVLAVG
VDNMFLMVAAVRRTSRTHTVHERMGECLADAAVSILITSSTDVLSFGVG
AITTIPAVQIFCVYTGVAIFFAFIYQITFFAACLALAMKHEASGRNSLF
LIEAVSAEKKTSLSTFQRLFNLGSVPDHSASHDVKQPLTSRFFGEWYAP
VLMHPVVRGIAMVWFVIYLLGASYGCSRIKEGLEPVNLLVEDSYAIPHY
RLLEKYFWKYGQQVQIVINNAPDLRNHTSRDRVHAMVLDFATSKHAIGM
ESVQFWLFEMERYYQKELEVQIIDSSFYGLLHHFLASKTNNPLAEDIYW
GPMPDDDNGTMVKSFRFILGMKDLVTTMDQTDATMSFREVAARWPEFNV
TTFMPIWMFTDQYIIIIPNTVQNIIIALLVMIVIAVLFIPQPMCSLWVA
LACASIDFGVIGYMTLWGVNLDAISMITIIMSIGFSVDYSAHIAYGYVV
SREDTAAGRVKEALSALGWPLSQGAMSTIIAVSVLADIPAYMIVTFFKT
VVLSISLGLLHGLVFLPVLLSIFVRGCCIIPSSPHGHPSAQKIEKQIRI
AAISSSPLDLRTVAPLRASSPISFPHRLEYTDESPTVHNRSKNSIKSE_
HLD

Fig. 18

GCTCGACTTTGGCCTGCAGGAGTAAGTGGCCCATTCCTAGGCTGTCCCC
AAAGATGTGTCCTCTGAAGAAGCTGGTCAGGTAGAGGGGGTGTCCACTA
TGGTTATAGTTGGGGAAGGAGATGCTGCCCAGGTCGAGCGTTTTGTCCA
CCTGCCAGGTGTACAGGAGTGGGTTGGGGGCACCACGCAGAGCATCTTG
TACCTCGCGCACACCTGCTGGTACTGGAGCTGGCTTGCGTTTCCCTGTG
CCACACGCAGATCCTGCACTGTACTGTCCAGTCAACTGACTTCAGCAAA
GATGTCTGGGTTCAGCAGCGAGTTGCTGTGTGAGGCCACCAGAATGGAG
GTGATATCGGCCTCAGTGCTCATCCTGGAGGTGGAGAAGTGGTCAGAGT
CCTCTGTGGTGAAATGGCTCTACACCAAAGTGGGAGACATTCAGAATCG
AGGGTCGA

Fig. 19

GATCCCCCTGTTTCACTTGGCATACGTTTTAATCCTACTCTTTGCTGTT
GTATCATGCTCCAGGTTGGACTGTATAAGAAACAAGATGTGTGTTGCAG
TCTTTGGAGTGTTTTCTGTTGCCATGTCAGTGGTGAGTGGTTTTGGCCT
GATGCTGCACCTTGGGGTCCCATTTGTGATTATAGTTGCAAATTCACCA
TTTCTTATTCTTGGTGAGTAAAAAAAATTACGAGGCTGTGCT

Fig. 20

TTTTATGGAATCAAGATTGACTTTCCAGAATGCCATGAAACCCGTTACC
CCTACAATCACAGAACCAATAGCAAAAGTCACCCACAAGGAACACAATG
GATAAGGGATTAACAATAAGGAAACAATGAACATAGCTGCTGATGCAAC
CAATACATTTCTAACAGTGTCTTCTAATATTGCAGCATACTGATCAAAA
TATATAAATGCCTGGTTATACACCATTAGGGGAATTTGACAGTCTTCAG
CTATGCGTCGTAATTAGAATAACAATATTTTCTTTTTGGCTGAGGAAGA
AACATCTGTTGTCTGAATGAAGCCCCGGGAAGAAATGATTTCATTTGAT
GAAGAAATATTAATATCATGCTGAAAATTTGGAAAATTGCTTAAAAAAT
CAGGAATATTGTTCATAAAAGTATTCTTCTCATTAGGATCTTGGCTGTT
ACCTTTTAAATATTGCACATATGCATCTAACCAAAACTCTGTAAGATTT
TTATCTACATAGACATTTTTTTCAAAAATTTTAGTACAGTTTTGCAGTT
TTTGCCTAACATCTTTATCCCAGTAGTCAACTTTTTTTAGTAACAATAA
CCATAACCCTGGGCCATAATCTGAAAAATA

Fig. 21

BLASTP - alignment of SEQ ID NO:1 against trembl|U88308|CEUC32E8_7

gene: "C32E8.8"; Caenorhabditis elegans cosmid C32E8. //:pironly|T25600|T25600 hypothetical protein C32E8.8 - Caenorhabditis elegans//:gp|U88308|1825729 gene: "C32E8.8"; Caenorhabditis elegans cosmid C32E8.

Pfam HMM search predict this Caenorhabditis elegans cosmid protein belong to the Patched family from residues 1 to 957 with score 1570.3 and E-value 0.

This hit is scoring at : 2e-38 (expectation value)
Alignment length (overlap) : 441
Identities : 28 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb

```
Q:    117 DCLEGLLSRTFQWLGWQVGAHPWIFLLAPLMLTAALGTGFLYLPKDEEEDLEEHYTPVGS
          DC:E  .:..F: LG: :  HP  F.: PL:.TAA:G.G.L:L  :. .D.   :TP:G:
H:      4 DCVERRAASLFRQLGFLICDHPLPFFVFPLLFTAAMGVGLLHL--NPLSDAVYLFTPLGA

          PAKAERRFVQGHFTTNDSYRFSASRRSTEANFVSLLVVSYSDS-LLDPATFAEVSKLDGA
          .:K.ER. :. .:...D: .: ..R. T::. :.:..:: :DS :LDP.   .V :LD
          QSKMERMSIHEKWPLTDN-NYIPGRAVTQSREIQVTALARNDSNILDPKFANAVYQLDKY

          VQDLRVAREKGSQIQYQQVCARYRALCVPPNPILYAWQVNKTLNLSSISFPAYNHG---R
          :Q.   . ..G....Y:.:C.:Y:   .P.N. ::.        LS.:    :NHG   .
          IQTRVRVLHNGHYYSYKNLCLQYKNGGCPSNKHVHI--------LSDL----HNHGFNIT
```

Fig. 21 (continued)

```
HPLYLTGFFGGYILGGSLG---------MGQLLLRAKAMRLLYYLKTEDPEYDVQSKQW-
:P.:  G  GGYI G.SLG           .:L..AKA. ::Y:LK ...E... S :W
YPYFRFGSEGGYI-GSSLGGVTVMKGENETDILASAKAWFMIYHLKFHPEEMSYISGEWE

--LTHLLDQFTNIKNILALKKIEVPGGVGLQGGQEKVVHFTSLSRQLEFEATSVTVIPVF
  L .:L.Q:.. . I                  .:.:F S : . E.:..: T:IP F
LELGRMLTQYPEDPYI-----------------SITYFHSQTLADELKRNADTLIPRF

HLAYILIILFA-------VTSCFRFDCIRNKMCVAAFGVISAFLAVVSGFGLLLHIGVPF
 ::..L:I:F:        :...F..D : :K .::..GV:SA :A:::G.G.L  :G:P:
IISITLLIVFSTLCSLSFIDGSFSIDWVLSKPILSILGVVSAGIAILTGVGFLSLMGMPY

VIIVANSPFLILGVGVDDMFIMISAWHKTNLAGDIRERMSNVYSKAAVSITITTITNILA
  IV. .PFL:L.VGVD:MF:M::A .:T: . .::.ERM.....:.AAVSI.IT: T::L:
NDIVGVMPFLVLAVGVDNMFLMVAAVRRTSRTHTVHERMGECLADAAVSILITSSTDVLS

LYTGIMSSFSIYGCFHVQEGL     534
. .G.::::.... .F V..G:
FGVGAITTIPAVQIFCVYTGV     409
```

Fig. 22

HMMPFAM - alignment of SEQ ID NO:1 against pfam|hmm|Patched

Patched family -

This hit is scoring at : -242.5
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:    166 ----DLEEHYTPVGSP-AKAERRFVQGHFTTnDSYRFSASRRSTEANFVSLLVVSYSD--
                .:   :TP .   : .ER . . . . .. :F ..  .. A.FV :.. : :.
H:      1 vlsskilytftpsDirksytergvrsedepl.verkFfpgdgkpiarfviitakarnsed

          ----SLLDPATFAEVSKLDGAVQD-LRVAREK---GSQIQYQQVCARYRALCVPPN----
              ::LD . . EV :::. : . :.:. .       .:.:::.:C:.:   : .P N
          ytrpNiLdikyLnevvqvndfiltnvtipvsgalkkitlsfkdlCssfcdincpvnqhyn

          ---P--ILYAWQVNktlNLSSISFPAYNHGRHPLYLTGFFGGYILGG-------------
              .  IL :   N   :  .:::P....   .:YL   FGG. L.
          gLdtlailrsnlhn...sridltYPtmtffgtkiylgpnfgGVkldpgeNdvekllDslt

          ------SLGMGQLLLRAKAMRLLYYLKTEDPEYDVQSKQWLTHLLDQFTNIKNILAlkKI
                : .    : ..KA: L.Y LK.:..E.: .SK:W ..L D...N  .  :  .I
          klklvdtkseegniksvkaivLyyrlkrdpeeveedskeWElsledyleneyksds..hi

          EVpggvglqggqeKVVHFTSLSRqlEFEATSVTVIPVFHLAYILIILFAVTSCFRFDC--
          EV            .:.. . L.   E. ...:T::P.F :.:.:::.F:V.:.. . .
          ev..........tifsdqvled..ElvrngltlvpffvvgfaiLvtFsvltsvvlavrr

          --------IRNKMCVAAFGVISAFLAVVSGFGLLLHIGVPFVIIVANSPFLILG-VGVDD
                  :.:K .:A..GV::..:A:VS.FGLL. :G. F  I:. .PFL:LG :GVDD
          sgslhidwvssKpiLAilgvltPlmAivsAfGlLfwlGfrFnsIlcVmPFLVLGAIGVDD
```

Fig. 22 (continued)

```
MFIMISAWHKTNLAGD------IRERMSNVYSKAAVSITITTITNILA----------
MF:M:.AW.:T. : .       :::RM..V. :A. SITIT::TN:L:
mFLMvhaWqrttrslssrarKvvkkRmgevlvEaGPSItITSlTnvLSFgIGaiTptPeI

--------------------LYTGIMS--------------------------------
                     .:..I::
qlFCiytavAiffdfiYQlTFFaAilaiagkyEmkgrhslflriTakavdaaetespkkl

-------------------------------------------------S----------

kkqrlksrgsakaaPvekegeknSkfseveeslesekssdssksheaeqktavvkffln

------------------------FSIYGCFHVQEGLDLRNLASDDSYITPYFNVEENYF
                        .:IYGC :::EGLD ..L. .DS :.....:.E.:.
iYcpfllnpkvrvcvllvyvvYlaiaiYGctnlkegLdPskLllkDSpLveilrlrekhv

SDYGPRVMVIVTKKVDYWDKDVRQKLENCTKIFEKNVY-VDK--NLTEFWLDAYVQYLKG
 .YG.:V.V:V... D. :.: R.:L .... FE.. Y :.K  N T:FWL  Y ::L..
wpyGsqvtVvVNNPPDLtnpenrdrllemvdeFEntpyalGkWEnsTkfWLrdYekFlss

------FMNNIVWEKL---------------SSCNYAINQTWLVkanasIPLYGPLNNK
      .:.:  E:                  . . Y : : :L       P :.  ...
sEeinleaseledeeeeidevdpvsvvvkdkqewifydllewFlk.....spgfshWqgd

MRKG-----PGGIVHT---RIL--VERLTCLLTSVLAVN------------LKVALWLSI
:  .     .  :..  ..    : L:. .... .:.              ..V.:: .
lvwdnktdketteikkFrFttgyHgkdlstwtdRtkllkewRgvAdeypddFNVTVFded

sppqlrpsselltqgprgtLAHISQLKSLSfPESFANFCLTADSKEARSQLLLPSGNC--
                  . .:.Q: S:  P.:.:.. T. .... . L.:P:  .
.................afFlDqilsil.pttiQsiiwtlicMavVCflFIpnpitvf
```

Fig. 22 (continued)

```
----------------------------PI--------------------YAETCWETK----
                           PI                     Y     : .
vitvsiaSIdiGVfGfLSLWGVdLDPisMitiIMSIGFSVDfsAHiaYhFyrskgsesrs

----CPSDSS-----------------------------------------------------
    ....
ikkiplktpdeRvadALealGWPviQAalSTiLcvlvLlFVpsYmvvvFfKTifLVVvlG

----------IWALQPP-------------SHSRSQGGPIP     783
          : :L  .              S::: :  . .
lLHGLifLPiiLslfvtssrifggdviirtsnakikkpseq     957
```

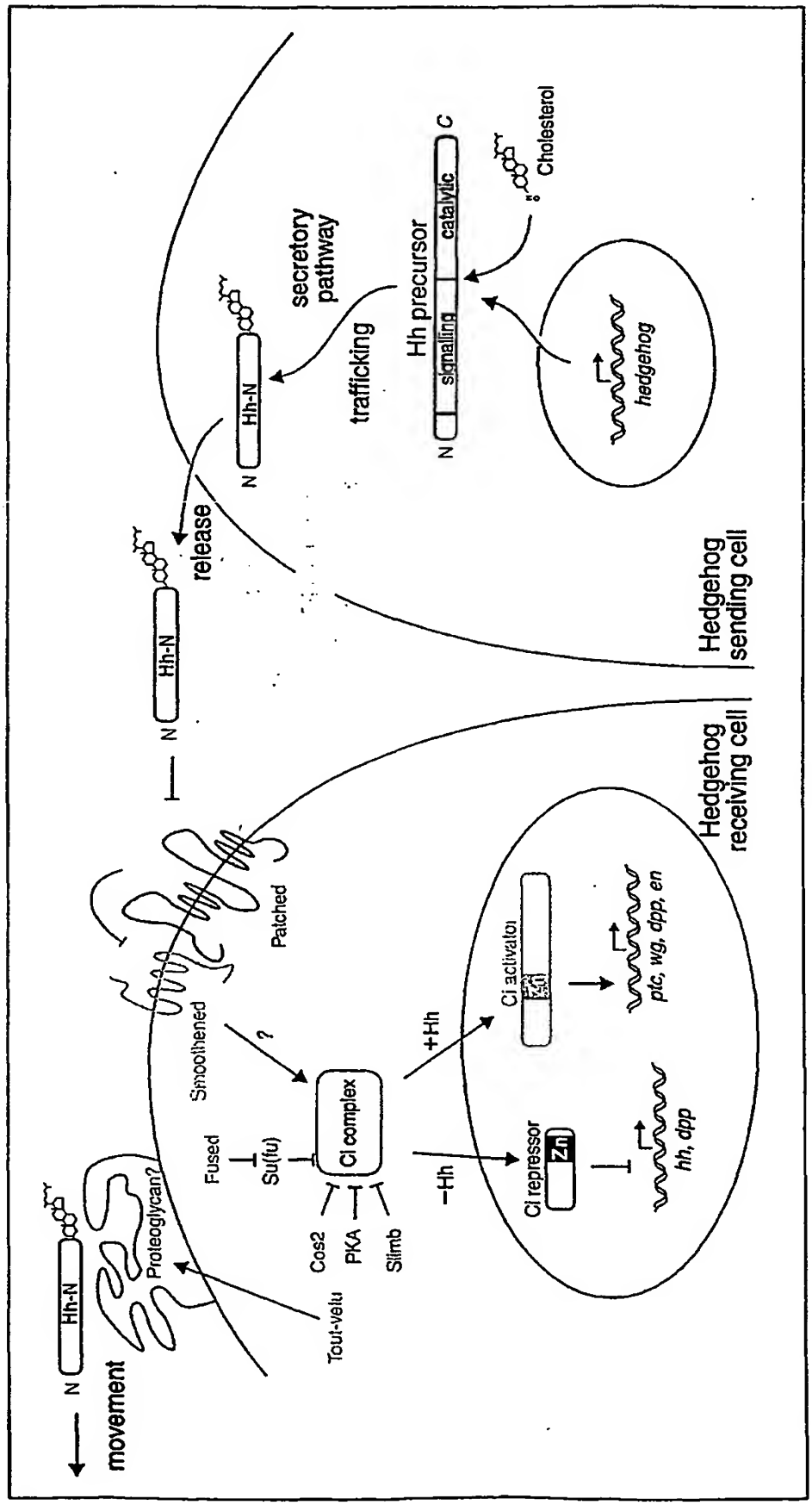
Hedgehog sending cell
Hedgehog receiving cell
secretory pathway
trafficking
release
movement
Hh precursor
signalling
catalytic
Cholesterol
hedgehog
Hh-N
Proteoglycan?
Tout-velu
Patched
Smoothened
Fused
Su(fu)
Cos2
PKA
Slimb
Ci complex
+Hh
−Hh
Ci activator
Ci repressor
Zn
ptc, wg, dpp, en
hh, dpp

Fig. 23

SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN PATCHED-LIKE PROTEIN

<130> LIO204 Foreign Countries

<150> US 60/245,565

<151> 2000-11-06

<150> US 60/245,564

<151> 2000-11-06

<150> US 60/245,572

<151> 2000-11-06

<160> 16

<170> PatentIn version 3.1

<210> 1

<211> 2355

<212> DNA

<213> Homo sapiens

<400> 1

atgccgtggg tggagcccaa gcccaggccg gggccggagc agaagcccaa gctcaccaaa 60

ccggactctg ccaccgggcc gcagtggtac caggaatctc aggaatcgga gtcggaaggc 120

aagcagccac ccccgggacc cctggcaccc ccgaaatccc ccgaaccctc aggacccctg 180

gcgtcggagc aggatgcacc cctgccagag ggggacgatg cacccccccg gccgtcgatg 240

ctggacgatg cacccgcct gccgctggag ctggacgatg caccctgcc ggaggaggaa 300

acccccgaac ccacggccat ctgcaggcac cggcaccgct gtcacaccga ctgcctagag 360

gggctgctgt cccgcacctt ccagtggctg gggtggcagg tgggcgcgca ccctggatc 420

ttcctgctgg cgcccttgat gctgacagcc gcgctgggca ccggcttcct gtacctaccc 480

aaggacgaag aggaagacct agaggagcat tacacccctg tggggagccc ggccaaggcg 540

gagcggcgct tcgtgcaggg ccatttcacc accaacgact cctaccgctt ctccgcctcc 600

aggaggagca ccgaagccaa tttcgtctcg cttctggtgg tctcctacag cgactcactg 660

ctggacccag ctacctttgc agaagtcagc aaactggacg gcgcggtgca ggatctgcgc 720

gtggcgcggg aaaagggaag ccagatccag taccagcagg tgtgcgcgag gtacagggcg 780

ctctgcgtgc ccccaaccc gatcctgtac gcctggcagg tgaacaaaac gctcaacctg 840

agcagcatct ccttccccgc ctacaaccac ggcaggcatc ccctctacct gaccggcttc 900

ttcggaggat acatcttggg gggcagccta ggaatgggcc agttactcct gcgggccaaa 960

gccatgcggc tgctgtacta cctgaagacc gaggaccctg agtacgacgt gcagagcaag 1020

cagtggctca cccatttgct cgatcaattt accaacatta agaacatctt ggccttgaaa 1080

aaaattgagg tacctggtgg tgtgggttta cagggaggcc aggagaaggt agtccacttt 1140

acatcgcttt ccagacaact ggaatttgag gcaacttctg tgactgtgat ccctgtgttt 1200

cacctggcat acattctcat cattctgttt gcagtcacat catgctttag cttctctatt 1260

tctgggggg aaatggttgt tggaattttg atggggattg tgttgaatct gcagattgtt 1320

ttgagtagga ctggcatttt aaaaatattg agtctgccaa cccaggaact aaggatcttt 1380

ccatttatcc aggtcttttt aagtttcttt taccaatgtt tgctggtttt cgatgaacat 1440

ggaactgata tccatccaat aagtttgttt tttagagact attttggccc ctttctcaca 1500

aggagtgagt ccaagtattt tgtagtcttt atatatgttt tgtacatcat aagcagtata 1560

tatgggtgtt tccatgtgca ggaaggttta gaccttcgaa atctggcaag tgacgattcc 1620

tacatcacac catattttaa cgtagaggag aattattttt cagattatgg tcccagggtt 1680

atggttattg ttactaaaaa agttgactac tgggataaag atgttaggca aaaactggaa 1740

aactgtacta aaatttttga aaaaaatgtc tatgtagata aaaatcttac agagttttgg 1800

ttagatgcat atgtgcaata tttaaaagga ttcatgaaca atattgtatg ggagaaactg 1860

agctcatgca actatgctat caatcagact tggctggtga aagccaatgc atctatcccc 1920

ttgtatgggc cactgaacaa taaaatgagg aaaggtccag gaggaattgt acatactagg 1980

atcttggtgg aaaggctcac ctgcctgctg acatcagttc tggcagtgaa cctgaaagtt 2040

gctctgtggc tcagcattag ccccccctcag ctgaggccca gctcagaact gctcacacaa 2100

ggacccagag ggacacttgc ccatatctca cagctcaaga gcctgagctt ccctgaaagt 2160

tttgccaact tctgtctcac agcagattcc aaggaggccc ggtctcagct cctgctgcca 2220

tcaggtaact gtcccatcta tgctgaaacg tgttgggaaa caaagtgccc ccctgattct 2280

tcgatatggg ctctccagcc tccatcccac agcagatccc aagggggtcc aattccagct 2340

ctggcctctc ttgtt 2355

<210> 2

<211> 785

<212> PRT

<213> Homo sapiens

<400> 2

Met Pro Trp Val Glu Pro Lys Pro Arg Pro Gly Pro Glu Gln Lys Pro
1 5 10 15

Lys Leu Thr Lys Pro Asp Ser Ala Thr Gly Pro Gln Trp Tyr Gln Glu
20 25 30

Ser Gln Glu Ser Glu Ser Glu Gly Lys Gln Pro Pro Pro Gly Pro Leu
35 40 45

Ala Pro Pro Lys Ser Pro Glu Pro Ser Gly Pro Leu Ala Ser Glu Gln
50 55 60

Asp Ala Pro Leu Pro Glu Gly Asp Asp Ala Pro Pro Arg Pro Ser Met
65 70 75 80

Leu Asp Asp Ala Pro Arg Leu Pro Leu Glu Leu Asp Asp Ala Pro Leu
85 90 95

Pro Glu Glu Glu Thr Pro Glu Pro Thr Ala Ile Cys Arg His Arg His
100 105 110

Arg Cys His Thr Asp Cys Leu Glu Gly Leu Leu Ser Arg Thr Phe Gln
115 120 125

Trp Leu Gly Trp Gln Val Gly Ala His Pro Trp Ile Phe Leu Leu Ala
130 135 140

```
Pro Leu Met Leu Thr Ala Ala Leu Gly Thr Gly Phe Leu Tyr Leu Pro
145                 150                 155                 160

Lys Asp Glu Glu Glu Asp Leu Glu Glu His Tyr Thr Pro Val Gly Ser
                165                 170                 175

Pro Ala Lys Ala Glu Arg Arg Phe Val Gln Gly His Phe Thr Thr Asn
            180                 185                 190

Asp Ser Tyr Arg Phe Ser Ala Ser Arg Arg Ser Thr Glu Ala Asn Phe
        195                 200                 205

Val Ser Leu Leu Val Val Ser Tyr Ser Asp Ser Leu Leu Asp Pro Ala
    210                 215                 220

Thr Phe Ala Glu Val Ser Lys Leu Asp Gly Ala Val Gln Asp Leu Arg
225                 230                 235                 240

Val Ala Arg Glu Lys Gly Ser Gln Ile Gln Tyr Gln Gln Val Cys Ala
                245                 250                 255

Arg Tyr Arg Ala Leu Cys Val Pro Pro Asn Pro Ile Leu Tyr Ala Trp
            260                 265                 270

Gln Val Asn Lys Thr Leu Asn Leu Ser Ser Ile Ser Phe Pro Ala Tyr
        275                 280                 285

Asn His Gly Arg His Pro Leu Tyr Leu Thr Gly Phe Phe Gly Gly Tyr
    290                 295                 300

Ile Leu Gly Gly Ser Leu Gly Met Gly Gln Leu Leu Leu Arg Ala Lys
305                 310                 315                 320

Ala Met Arg Leu Leu Tyr Tyr Leu Lys Thr Glu Asp Pro Glu Tyr Asp
                325                 330                 335

Val Gln Ser Lys Gln Trp Leu Thr His Leu Leu Asp Gln Phe Thr Asn
            340                 345                 350

Ile Lys Asn Ile Leu Ala Leu Lys Lys Ile Glu Val Pro Gly Gly Val
        355                 360                 365

Gly Leu Gln Gly Gly Gln Glu Lys Val Val His Phe Thr Ser Leu Ser
    370                 375                 380
```

Arg Gln Leu Glu Phe Glu Ala Thr Ser Val Thr Val Ile Pro Val Phe
385 390 395 400

His Leu Ala Tyr Ile Leu Ile Ile Leu Phe Ala Val Thr Ser Cys Phe
405 410 415

Arg Phe Asp Cys Ile Arg Asn Lys Met Cys Val Ala Ala Phe Gly Val
420 425 430

Ile Ser Ala Phe Leu Ala Val Val Ser Gly Phe Gly Leu Leu Leu His
435 440 445

Ile Gly Val Pro Phe Val Ile Ile Val Ala Asn Ser Pro Phe Leu Ile
450 455 460

Leu Gly Val Gly Val Asp Asp Met Phe Ile Met Ile Ser Ala Trp His
465 470 475 480

Lys Thr Asn Leu Ala Gly Asp Ile Arg Glu Arg Met Ser Asn Val Tyr
485 490 495

Ser Lys Ala Ala Val Ser Ile Thr Ile Thr Thr Ile Thr Asn Ile Leu
500 505 510

Ala Leu Tyr Thr Gly Ile Met Ser Ser Phe Ser Ile Tyr Gly Cys Phe
515 520 525

His Val Gln Glu Gly Leu Asp Leu Arg Asn Leu Ala Ser Asp Asp Ser
530 535 540

Tyr Ile Thr Pro Tyr Phe Asn Val Glu Glu Asn Tyr Phe Ser Asp Tyr
545 550 555 560

Gly Pro Arg Val Met Val Ile Val Thr Lys Lys Val Asp Tyr Trp Asp
565 570 575

Lys Asp Val Arg Gln Lys Leu Glu Asn Cys Thr Lys Ile Phe Glu Lys
580 585 590

Asn Val Tyr Val Asp Lys Asn Leu Thr Glu Phe Trp Leu Asp Ala Tyr
595 600 605

Val Gln Tyr Leu Lys Gly Phe Met Asn Asn Ile Val Trp Glu Lys Leu
610 615 620

```
Ser Ser Cys Asn Tyr Ala Ile Asn Gln Thr Trp Leu Val Lys Ala Asn
625                 630                 635                 640

Ala Ser Ile Pro Leu Tyr Gly Pro Leu Asn Asn Lys Met Arg Lys Gly
                645                 650                 655

Pro Gly Gly Ile Val His Thr Arg Ile Leu Val Glu Arg Leu Thr Cys
            660                 665                 670

Leu Leu Thr Ser Val Leu Ala Val Asn Leu Lys Val Ala Leu Trp Leu
        675                 680                 685

Ser Ile Ser Pro Pro Gln Leu Arg Pro Ser Ser Glu Leu Leu Thr Gln
    690                 695                 700

Gly Pro Arg Gly Thr Leu Ala His Ile Ser Gln Leu Lys Ser Leu Ser
705                 710                 715                 720

Phe Pro Glu Ser Phe Ala Asn Phe Cys Leu Thr Ala Asp Ser Lys Glu
                725                 730                 735

Ala Arg Ser Gln Leu Leu Leu Pro Ser Gly Asn Cys Pro Ile Tyr Ala
            740                 745                 750

Glu Thr Cys Trp Glu Thr Lys Cys Pro Ser Asp Ser Ser Ile Trp Ala
        755                 760                 765

Leu Gln Pro Pro Ser His Ser Arg Ser Gln Gly Gly Pro Ile Pro Ala
    770                 775                 780

Leu
785
```

<210> 3

<211> 933

<212> PRT

<213> Caenorhabditis elegans

<400> 3

```
Met Ala Trp Asp Cys Val Glu Arg Arg Ala Ala Ser Leu Phe Arg Gln
1               5                  10                  15

Leu Gly Phe Leu Ile Cys Asp His Pro Leu Pro Phe Phe Val Phe Pro
            20                  25                  30

Leu Leu Phe Thr Ala Ala Met Gly Val Gly Leu Leu His Leu Asn Pro
        35                  40                  45

Leu Ser Asp Ala Val Tyr Leu Phe Thr Pro Leu Gly Ala Gln Ser Lys
    50                  55                  60

Met Glu Arg Met Ser Ile His Glu Lys Trp Pro Leu Thr Asp Asn Asn
65                  70                  75                  80

Tyr Ile Pro Gly Arg Ala Val Thr Gln Ser Arg Glu Ile Gln Val Thr
                85                  90                  95

Ala Leu Ala Arg Asn Asp Ser Asn Ile Leu Asp Pro Lys Phe Ala Asn
            100                 105                 110

Ala Val Tyr Gln Leu Asp Lys Tyr Ile Gln Thr Arg Val Arg Val Leu
        115                 120                 125

His Asn Gly His Tyr Tyr Ser Tyr Lys Asn Leu Cys Leu Gln Tyr Lys
    130                 135                 140

Asn Gly Gly Cys Pro Ser Asn Lys His Val His Ile Leu Ser Asp Leu
145                 150                 155                 160

His Asn His Gly Phe Asn Ile Thr Tyr Pro Tyr Phe Arg Phe Gly Ser
                165                 170                 175

Glu Gly Gly Tyr Ile Gly Ser Ser Leu Gly Gly Val Thr Val Met Lys
            180                 185                 190

Gly Glu Asn Glu Thr Asp Ile Leu Ala Ser Ala Lys Ala Trp Phe Met
        195                 200                 205

Ile Tyr His Leu Lys Phe His Pro Glu Glu Met Ser Tyr Ile Ser Gly
    210                 215                 220

Glu Trp Glu Leu Glu Leu Gly Arg Met Leu Thr Gln Tyr Pro Glu Asp
225                 230                 235                 240
```

```
Pro Tyr Ile Ser Ile Thr Tyr Phe His Ser Gln Thr Leu Ala Asp Glu
                245                 250                 255

Leu Lys Arg Asn Ala Asp Thr Leu Ile Pro Arg Phe Ile Ile Ser Ile
            260                 265                 270

Thr Leu Leu Ile Val Phe Ser Thr Leu Cys Ser Leu Ser Phe Ile Asp
        275                 280                 285

Gly Ser Phe Ser Ile Asp Trp Val Leu Ser Lys Pro Ile Leu Ser Ile
    290                 295                 300

Leu Gly Val Val Ser Ala Gly Ile Ala Ile Leu Thr Gly Val Gly Phe
305                 310                 315                 320

Leu Ser Leu Met Gly Met Pro Tyr Asn Asp Ile Val Gly Val Met Pro
                325                 330                 335

Phe Leu Val Leu Ala Val Gly Val Asp Asn Met Phe Leu Met Val Ala
            340                 345                 350

Ala Val Arg Arg Thr Ser Arg Thr His Thr Val His Glu Arg Met Gly
        355                 360                 365

Glu Cys Leu Ala Asp Ala Ala Val Ser Ile Leu Ile Thr Ser Ser Thr
    370                 375                 380

Asp Val Leu Ser Phe Gly Val Gly Ala Ile Thr Thr Ile Pro Ala Val
385                 390                 395                 400

Gln Ile Phe Cys Val Tyr Thr Gly Val Ala Ile Phe Phe Ala Phe Ile
                405                 410                 415

Tyr Gln Ile Thr Phe Phe Ala Ala Cys Leu Ala Leu Ala Met Lys His
            420                 425                 430

Glu Ala Ser Gly Arg Asn Ser Leu Phe Leu Ile Glu Ala Val Ser Ala
        435                 440                 445

Glu Lys Lys Thr Ser Leu Ser Thr Phe Gln Arg Leu Phe Asn Leu Gly
    450                 455                 460

Ser Val Pro Asp His Ser Ala Ser His Asp Val Lys Gln Pro Leu Thr
465                 470                 475                 480
```

Ser Arg Phe Phe Gly Glu Trp Tyr Ala Pro Val Leu Met His Pro Val
485 490 495

Val Arg Gly Ile Ala Met Val Trp Phe Val Ile Tyr Leu Leu Gly Ala
500 505 510

Ser Tyr Gly Cys Ser Arg Ile Lys Glu Gly Leu Glu Pro Val Asn Leu
515 520 525

Leu Val Glu Asp Ser Tyr Ala Ile Pro His Tyr Arg Leu Leu Glu Lys
530 535 540

Tyr Phe Trp Lys Tyr Gly Gln Gln Val Gln Ile Val Ile Asn Asn Ala
545 550 555 560

Pro Asp Leu Arg Asn His Thr Ser Arg Asp Arg Val His Ala Met Val
565 570 575

Leu Asp Phe Ala Thr Ser Lys His Ala Ile Gly Met Glu Ser Val Gln
580 585 590

Phe Trp Leu Phe Glu Met Glu Arg Tyr Tyr Gln Lys Glu Leu Glu Val
595 600 605

Gln Ile Ile Asp Ser Ser Phe Tyr Gly Leu Leu His His Phe Leu Ala
610 615 620

Ser Lys Thr Asn Asn Pro Leu Ala Glu Asp Ile Tyr Trp Gly Pro Met
625 630 635 640

Pro Asp Asp Asp Asn Gly Thr Met Val Lys Ser Phe Arg Phe Ile Leu
645 650 655

Gly Met Lys Asp Leu Val Thr Thr Met Asp Gln Thr Asp Ala Thr Met
660 665 670

Ser Phe Arg Glu Val Ala Ala Arg Trp Pro Glu Phe Asn Val Thr Thr
675 680 685

Phe Met Pro Ile Trp Met Phe Thr Asp Gln Tyr Ile Ile Ile Ile Pro
690 695 700

Asn Thr Val Gln Asn Ile Ile Ile Ala Leu Leu Val Met Ile Val Ile
705 710 715 720

Ala Val Leu Phe Ile Pro Gln Pro Met Cys Ser Leu Trp Val Ala Leu
725 730 735

Ala Cys Ala Ser Ile Asp Phe Gly Val Ile Gly Tyr Met Thr Leu Trp
740 745 750

Gly Val Asn Leu Asp Ala Ile Ser Met Ile Thr Ile Ile Met Ser Ile
755 760 765

Gly Phe Ser Val Asp Tyr Ser Ala His Ile Ala Tyr Gly Tyr Val Val
770 775 780

Ser Arg Glu Asp Thr Ala Ala Gly Arg Val Lys Glu Ala Leu Ser Ala
785 790 795 800

Leu Gly Trp Pro Leu Ser Gln Gly Ala Met Ser Thr Ile Ile Ala Val
805 810 815

Ser Val Leu Ala Asp Ile Pro Ala Tyr Met Ile Val Thr Phe Phe Lys
820 825 830

Thr Val Val Leu Ser Ile Ser Leu Gly Leu Leu His Gly Leu Val Phe
835 840 845

Leu Pro Val Leu Leu Ser Ile Phe Val Arg Gly Cys Cys Ile Ile Pro
850 855 860

Ser Ser Pro His Gly His Pro Ser Ala Gln Lys Ile Glu Lys Gln Ile
865 870 875 880

Arg Ile Ala Ala Ile Ser Ser Ser Pro Leu Asp Leu Arg Thr Val Ala
885 890 895

Pro Leu Arg Ala Ser Ser Pro Ile Ser Phe Pro His Arg Leu Glu Tyr
900 905 910

Thr Asp Glu Ser Pro Thr Val His Asn Arg Ser Lys Asn Ser Ile Lys
915 920 925

Ser Glu His Leu Asp
930

<210> 4

<211> 449

<212> DNA

<213> Homo sapiens

<400> 4
gctcgacttt ggcctgcagg agtaagtggc ccattcctag gctgtcccca aagatgtgtc 60

ctctgaagaa gctggtcagg tagagggggt gtccactatg gttatagttg gggaaggaga 120

tgctgcccag gtcgagcgtt ttgtccacct gccaggtgta caggagtggg ttgggggcac 180

cacgcagagc atcttgtacc tcgcgcacac ctgctggtac tggagctggc ttgcgtttcc 240

ctgtgccaca cgcagatcct gcactgtact gtccagtcaa ctgacttcag caäägatgtc 300

tgggttcagc agcgagttgc tgtgtgaggc caccagaatg gaggtgatat cggcctcagt 360

gctcatcctg gaggtggaga agtggtcaga gtcctctgtg gtgaaatggc tctacaccaa 420

agtgggagac attcagaatc gagggtcga 449

<210> 5

<211> 238

<212> DNA

<213> Homo sapiens

<400> 5
gatccccctg tttcacttgg catacgtttt aatcctactc tttgctgttg tatcatgctc 60

caggttggac tgtataagaa acaagatgtg tgttgcagtc tttggagtgt tttctgttgc 120

catgtcagtg gtgagtggtt ttggcctgat gctgcacctt ggggtcccat ttgtgattat 180

agttgcaaat tcaccatttc ttattcttgg tgagtaaaaa aaattacgag gctgtgct 238

<210> 6

<211> 618

<212> DNA

<213> Homo sapiens

<400> 6
ttttatggaa tcaagattga ctttccagaa tgccatgaaa cccgttaccc ctacaatcac 60
agaaccaata gcaaaagtca cccacaagga acacaatgga taagggatta acaataagga 120
aacaatgaac atagctgctg atgcaaccaa tacatttcta acagtgtctt ctaatattgc 180
agcatactga tcaaaatata taaatgcctg gttatacacc attaggggaa tttgacagtc 240
ttcagctatg cgtcgtaatt agaataacaa tattttcttt ttggctgagg aagaaacatc 300
tgttgtctga atgaagcccc gggaagaaat gatttcattt gatgaagaaa tattaatatc 360
atgctgaaaa tttggaaaat tgcttaaaaa atcaggaata ttgttcataa aagtattctt 420
ctcattagga tcttggctgt taccttttaa atattgcaca tatgcatcta accaaaactc 480
tgtaagattt ttatctacat agacattttt ttcaaaaatt ttagtacagt tttgcagttt 540
ttgcctaaca tctttatccc agtagtcaac tttttttagt aacaataacc atäaccctgg 600
gccataatct gaaaaata 618

<210> 7
<211> 2379
<212> DNA
<213> Homo sapiens

<400> 7
atgaaatcac aagccacaca atcactcagt gaaactagcg ggactgctaa ggtcctttcc 60
catttactga gagagccaca ggaaccgaga ggagaatatc agcatcctct tatcttgagc 120
cacactgcac agtcccttct cactccactg atgtcaggtc ctccagcagc cttcagcact 180
ttgtgggatt ggcagagcag aagcagtcac cggtctttgc tcacactcac agcttcaggg 240
catgaacaca catcacactc tctcaagaaa cctctcacca tgcactggat gataacactg 300
aaacaggatt cagattctgc agctcggaaa tctttcagca tctttaccaa ttcagtcaat 360
gtcatcactg accatggggt tgccaaggct caaaacctgg acatcatcct ttactccttt 420
cctccaactt cattgtactg tgttacacaa ggtcctatat ggaaaaagaa tccacagtgg 480
aattggaagt ccaggctccc tgctcaacct ttgccagcag cacaggggaa cagagggcat 540
tctgttcaga aaaaaatcca gctgtggaat gtttttattt catcttcatt gggaagaaag 600
gccacaaaac attcagggat gcccccttaaa acttgtcctt tgaagcatta tgcttttatc 660
aagcatctgt gctacagctt tgaagatttt tctcttgaat catatttagt agaaatcaaa 720
gctgtgcata atctaagttt gcaaagtcat ggaactaaag gagtgtttga gcttctgtcc 780

ggatggcgga gaaccaaaga gaacttgccc ttcaaagaca ggatagcaga tgcctattct 840

gatgtgatgg tcacctatac catgaccagc tccctgtact tcatcacttt tggcatgggt 900

gccagcccat tcacaaacat agaggctgtg aaggtcttct gtcaaaacat gtgtgtctct 960

attctgttga actacttcta cattttctcc ttctttggct cctgtctggt ctttgctggc 1020

caactagagc aaaaccgcta ccacagcatc ttttgctgta agatcccttc tgcagaatac 1080

ctggatcgca aacctgtgtg gttccagaca gtgatgagtg atgggcatca acagacgtcc 1140

catcatgaga cgaacccctа ccagcaccac ttcattcagc acttcctccg tgaacattat 1200

aatgaatgga ttaccaatat atatgtgaag ccatttgttg tcatcctcta tctcatttat 1260

gcctccttct ccttcatggg gtgcttacag atcagtgacg gagccaacat catcaatcta 1320

ctagccagtg attcgccaag tgtttcctat gccatggttc agcagaaata tttcagcaac 1380

tatagccctg tgataggatt ctacgtctat gagcccctag agtactggaa cagcagcgtc 1440

caggatgacc taagaagact ctgtagtgga ttcactgcag tgtcctgggt ggagcagtac 1500

taccagttcc tgaaagtcag caacgtcagt gccaataaca aaagtgactt catcagtgtc 1560

ctgcaaagct catttttaaa aaagccagaa ttccagcatt ttcgaaatga tatcatcttc 1620

tccaaggcag gggatgaaag caatatcatt gcttctcgct tgtatctggt ggccaggact 1680

agcagagaca agcagaaaga aatcacagaa gtgttggaaa agctgaggcc cctatccctc 1740

tcaaagagca tccgattcat cgtgttcaac ccctcctttg tcttcatgga ccattacagc 1800

ttgtctgtca cagtgcctgt tctgattgca ggctttggtg ttctcctggt gttaatcctg 1860

actttttтcc tagtgatcca ccctctggga aacttctggc taattcttag cgtcacctca 1920

attgagctgg gcgttctggg cttaatgaca ttatggaacg tcgacatgga ttgcatttct 1980

atcttgtgcc ttatctacac cttgaatttc gccattgacc actgtgcacc actgcttttc 2040

acatttgtat tagcaactga gcacacccga acacaatgta taaaaagctc cttgcaagac 2100

catgggacag ccattttgca aaatgttact tcttttctta ttgggttagt cccccttcta 2160

tttgtgcctt cgaacctgac cttcacactg ttcaaatgct tgctgctcac tgggggttgc 2220

acacttctgc actgttttgt tattttacct gtgttcctaa cgtttttccc cccttccaaa 2280

aagcaccaca agaaaaagaa acgtgccaag cgaaaggaga gagaggaaat tgaatgcata 2340

gaaattcaag agaacccgga tcacgtcacc acagtatga 2379

-<210> 8

<211> 792

<212> PRT

<213> Homo sapiens

<400> 8

```
Met Lys Ser Gln Ala Thr Gln Ser Leu Ser Glu Thr Ser Gly Thr Ala
1               5                  10                  15

Lys Val Leu Ser His Leu Leu Arg Glu Pro Gln Glu Pro Arg Gly Glu
            20                  25                  30

Tyr Gln His Pro Leu Ile Leu Ser His Thr Ala Gln Ser Leu Leu Thr
        35                  40                  45

Pro Leu Met Ser Gly Pro Pro Ala Ala Phe Ser Thr Leu Trp Asp Trp
    50                  55                  60

Gln Ser Arg Ser Ser His Arg Ser Leu Leu Thr Leu Thr Ala Ser Gly
65                  70                  75                  80

His Glu His Thr Ser His Ser Leu Lys Lys Pro Leu Thr Met His Trp
                85                  90                  95

Met Ile Thr Leu Lys Gln Asp Ser Asp Ser Ala Ala Arg Lys Ser Phe
            100                 105                 110

Ser Ile Phe Thr Asn Ser Val Asn Val Ile Thr Asp His Gly Val Ala
        115                 120                 125

Lys Ala Gln Asn Leu Asp Ile Ile Leu Tyr Ser Phe Pro Pro Thr Ser
    130                 135                 140

Leu Tyr Cys Val Thr Gln Gly Pro Ile Trp Lys Lys Asn Pro Gln Trp
145                 150                 155                 160

Asn Trp Lys Ser Arg Leu Pro Ala Gln Pro Leu Pro Ala Ala Gln Gly
                165                 170                 175

Asn Arg Gly His Ser Val Gln Lys Lys Ile Gln Leu Trp Asn Val Phe
            180                 185                 190
```

Ile Ser Ser Ser Leu Gly Arg Lys Ala Thr Lys His Ser Gly Met Pro
195 200 205

Leu Lys Thr Cys Pro Leu Lys His Tyr Ala Phe Ile Lys His Leu Cys
210 215 220

Tyr Ser Phe Glu Asp Phe Ser Leu Glu Ser Tyr Leu Val Glu Ile Lys
225 230 235 240

Ala Val His Asn Leu Ser Leu Gln Ser His Gly Thr Lys Gly Val Phe
245 250 255

Glu Leu Leu Ser Gly Trp Arg Arg Thr Lys Glu Asn Leu Pro Phe Lys
260 265 270

Asp Arg Ile Ala Asp Ala Tyr Ser Asp Val Met Val Thr Tyr Thr Met
275 280 285

Thr Ser Ser Leu Tyr Phe Ile Thr Phe Gly Met Gly Ala Ser Pro Phe
290 295 300

Thr Asn Ile Glu Ala Val Lys Val Phe Cys Gln Asn Met Cys Val Ser
305 310 315 320

Ile Leu Leu Asn Tyr Phe Tyr Ile Phe Ser Phe Phe Gly Ser Cys Leu
325 330 335

Val Phe Ala Gly Gln Leu Glu Gln Asn Arg Tyr His Ser Ile Phe Cys
340 345 350

Cys Lys Ile Pro Ser Ala Glu Tyr Leu Asp Arg Lys Pro Val Trp Phe
355 360 365

Gln Thr Val Met Ser Asp Gly His Gln Gln Thr Ser His His Glu Thr
370 375 380

Asn Pro Tyr Gln His His Phe Ile Gln His Phe Leu Arg Glu His Tyr
385 390 395 400

Asn Glu Trp Ile Thr Asn Ile Tyr Val Lys Pro Phe Val Val Ile Leu
405 410 415

Tyr Leu Ile Tyr Ala Ser Phe Ser Phe Met Gly Cys Leu Gln Ile Ser
420 425 430

```
Asp Gly Ala Asn Ile Ile Asn Leu Leu Ala Ser Asp Ser Pro Ser Val
        435                 440                 445

Ser Tyr Ala Met Val Gln Gln Lys Tyr Phe Ser Asn Tyr Ser Pro Val
    450                 455                 460

Ile Gly Phe Tyr Val Tyr Glu Pro Leu Glu Tyr Trp Asn Ser Ser Val
465                 470                 475                 480

Gln Asp Asp Leu Arg Arg Leu Cys Ser Gly Phe Thr Ala Val Ser Trp
                485                 490                 495

Val Glu Gln Tyr Tyr Gln Phe Leu Lys Val Ser Asn Val Ser Ala Asn
            500                 505                 510

Asn Lys Ser Asp Phe Ile Ser Val Leu Gln Ser Ser Phe Leu Lys Lys
        515                 520                 525

Pro Glu Phe Gln His Phe Arg Asn Asp Ile Ile Phe Ser Lys Ala Gly
    530                 535                 540

Asp Glu Ser Asn Ile Ile Ala Ser Arg Leu Tyr Leu Val Ala Arg Thr
545                 550                 555                 560

Ser Arg Asp Lys Gln Lys Glu Ile Thr Glu Val Leu Glu Lys Leu Arg
                565                 570                 575

Pro Leu Ser Leu Ser Lys Ser Ile Arg Phe Ile Val Phe Asn Pro Ser
            580                 585                 590

Phe Val Phe Met Asp His Tyr Ser Leu Ser Val Thr Val Pro Val Leu
        595                 600                 605

Ile Ala Gly Phe Gly Val Leu Leu Val Leu Ile Leu Thr Phe Phe Leu
    610                 615                 620

Val Ile His Pro Leu Gly Asn Phe Trp Leu Ile Leu Ser Val Thr Ser
625                 630                 635                 640

Ile Glu Leu Gly Val Leu Gly Leu Met Thr Leu Trp Asn Val Asp Met
                645                 650                 655

Asp Cys Ile Ser Ile Leu Cys Leu Ile Tyr Thr Leu Asn Phe Ala Ile
            660                 665                 670
```

Asp His Cys Ala Pro Leu Leu Phe Thr Phe Val Leu Ala Thr Glu His
675 680 685

Thr Arg Thr Gln Cys Ile Lys Ser Ser Leu Gln Asp His Gly Thr Ala
690 695 700

Ile Leu Gln Asn Val Thr Ser Phe Leu Ile Gly Leu Val Pro Leu Leu
705 710 715 720

Phe Val Pro Ser Asn Leu Thr Phe Thr Leu Phe Lys Cys Leu Leu Leu
725 730 735

Thr Gly Gly Cys Thr Leu Leu His Cys Phe Val Ile Leu Pro Val Phe
740 745 750

Leu Thr Phe Phe Pro Pro Ser Lys Lys His His Lys Lys Lys Lys Arg
755 760 765

Ala Lys Arg Lys Glu Arg Glu Glu Ile Glu Cys Ile Glu Ile Gln Glu
770 775 780

Asn Pro Asp His Val Thr Thr Val
785 790

<210> 9

<211> 1061

<212> PRT

<213> Drosophila melanogaster

<400> 9

Met Ile Arg Arg Glu Val Phe Gln Glu Leu Arg Gln Leu Asp Asn Ile
1 5 10 15

Ile Gln Asn Ala Thr Thr Thr Tyr Asp Gly Asp Thr Tyr Thr Tyr Lys
20 25 30

Asp Asn Cys Ala Arg Trp Glu Asn Glu Cys Phe Glu Asn Asp Ile Leu
35 40 45

Asn Leu Asp Ala Leu Met Asp Asp Ile Glu Ala Gly Gln Leu Asn Leu
50 55 60

```
Thr Phe Pro Phe Met Phe Asn Pro Val Thr Trp Asp Ala His Leu Phe
65                  70                  75                  80

Pro Val Phe Phe Gly Gly Thr Lys Leu Thr Glu Asp Asn Tyr Val Ile
                85                  90                  95

Ser Val Pro Ala Ile Gln Leu Val Tyr Phe Val Thr Ala Asp Thr Lys
            100                 105                 110

Arg Gln Asp Ala Lys Gly Ala Glu Trp Glu Glu Thr Phe Leu Arg Val
        115                 120                 125

Val Gly Asn Ala Glu Asn Ser Gly Gln Phe Lys His Ile Ser Val Ser
    130                 135                 140

Tyr Phe Ala Ser Arg Thr Leu Asp His Glu Leu Glu Lys Asn Thr Lys
145                 150                 155                 160

Thr Val Val Pro Tyr Phe Ser Ser Thr Phe Leu Leu Met Gly Leu Phe
                165                 170                 175

Ser Ile Ile Thr Cys Met Met Gly Asp Ala Val Arg Ser Lys Pro Phe
            180                 185                 190

Leu Gly Leu Met Gly Asn Val Ser Ala Ile Met Ala Thr Leu Ala Ala
        195                 200                 205

Phe Gly Leu Ala Met Tyr Cys Gly Ile Glu Phe Ile Gly Ile Asn Leu
    210                 215                 220

Ala Ala Pro Phe Leu Met Ile Gly Ile Gly Ile Asp Asp Thr Phe Val
225                 230                 235                 240

Met Leu Ala Gly Trp Arg Arg Thr Lys Ala Lys Met Pro Val Ala Glu
                245                 250                 255

Arg Met Gly Leu Met Met Ser Glu Ala Ala Val Ser Ile Thr Ile Thr
            260                 265                 270

Ser Val Thr Asp Phe Ile Ser Phe Leu Ile Gly Ile Ile Ser Pro Phe
        275                 280                 285

Arg Ser Val Arg Ile Phe Cys Thr Tyr Ser Val Phe Ala Val Cys Phe
    290                 295                 300
```

```
Thr Phe Leu Trp His Ile Thr Phe Phe Ala Ala Cys Met Ala Ile Ser
305                 310                 315                 320

Gly Tyr Arg Glu Arg Lys Asn Leu His Ser Ile Phe Gly Cys Arg Val
                325                 330                 335

Gln Pro Met Ser Val Ala Ile Lys Glu Lys Arg Asn Phe Leu Tyr Lys
            340                 345                 350

Ala Ile Met Ala Gly Gly Ile Asp Ala Asn Asp Pro Asp Asn Pro Ile
        355                 360                 365

Asp Asn Lys Asp His Met Leu Met Ala Phe Phe Lys Asp Lys Met Ala
    370                 375                 380

Ala Val Ile Asn Asn Lys Trp Cys Lys Ala Ile Ile Ile Leu Ala Phe
385                 390                 395                 400

Ala Ser Tyr Leu Val Gly Ala Cys Tyr Gly Ile Thr Gln Ile Lys Glu
                405                 410                 415

Gly Leu Glu Arg Arg Lys Leu Ser Arg Glu Asp Ser Tyr Ser Val Glu
            420                 425                 430

Phe Phe Asp Arg Glu Asp Asp Tyr Tyr Arg Glu Phe Pro Tyr Arg Met
        435                 440                 445

Gln Val Ile Ile Ala Gly Pro Leu Asn Tyr Ser Asp Pro Leu Val Gln
    450                 455                 460

Glu Gln Val Glu Asn Leu Thr Ser Thr Leu Glu His Thr Ser Tyr Val
465                 470                 475                 480

Thr Ser Arg Arg Tyr Thr Glu Ser Trp Leu Arg Ser Phe Leu Ser Phe
                485                 490                 495

Leu Glu Arg Asn Asn Glu Leu Leu Asn Val Thr Val Asp Asp Glu Gln
            500                 505                 510

Thr Phe Ile Asp Ala Val Lys Glu His Trp Leu Phe Pro Gly Asn Pro
        515                 520                 525

Phe Ser Leu Asp Val Arg Phe Asn Glu Asp Glu Thr Gln Ile Ile Ala
    530                 535                 540
```

Ser Arg Phe Leu Ile Gln Ala Val Asn Ile Thr Asp Thr Asn His Glu
545 550 555 560

Lys Glu Met Val Arg Asp Leu Arg Gln Ile Cys Lys Asp Ser Pro Leu
565 570 575

Asn Ala Ser Ile Phe His Pro Tyr Phe Val Phe Phe Asp Gln Phe Glu
580 585 590

Leu Val Arg Pro Val Ser Leu Gln Ala Met Val Ile Gly Ala Ile Ile
595 600 605

Met Met Ile Ile Ser Phe Val Phe Ile Pro Asn Ile Leu Cys Ser Leu
610 615 620

Trp Val Ala Phe Ser Val Ile Ser Ile Glu Leu Gly Val Ala Gly Tyr
625 630 635 640

Met Ala Leu Trp Asp Val Asn Leu Asp Ser Ile Ser Met Ile Asn Leu
645 650 655

Ile Met Cys Ile Gly Phe Ser Val Asp Phe Thr Ala His Ile Cys Tyr
660 665 670

Thr Tyr Met Ser Ser Lys Lys Arg Ser Pro Lys Ala Arg Val Arg Glu
675 680 685

Ala Leu His Ser Leu Gly Leu Pro Ile Ile Gln Gly Ser Ser Ser Thr
690 695 700

Ile Leu Gly Ile Val Ala Leu Leu Leu Ala Gln Ser Tyr Ile Phe Leu
705 710 715 720

Val Phe Phe Lys Met Val Phe Leu Val Ile Phe Phe Gly Ala Met His
725 730 735

Gly Leu Phe Leu Leu Pro Val Leu Leu Ser Leu Phe Gly Pro Gly Ser
740 745 750

Cys Leu Thr Trp Thr Gly Lys Asp Asp Gly Ser Asp Ala Glu Val Asp
755 760 765

Asp Gly Leu Asp Asp Arg Gln Leu Glu Lys Pro Phe Ser Gln Ser Tyr
770 775 780

```
Tyr Met Gln Tyr Pro Ser Ile Gly Ile Asn Gly Pro Tyr Gly Ser Lys
785                 790                 795                 800

Gly Phe Leu Gly Ala Pro Tyr Lys Ala Tyr Gly Val Asp Glu Lys Asp
                805                 810                 815

Leu Gly Leu Gly Thr Ser Gly Glu Asp Ser Ser Glu Ser Ser Ser Ser
            820                 825                 830

Arg Ser Gln His Arg Gln Gln Ala Ala Ala Thr Glu Glu Glu Val Val
        835                 840                 845

Val Arg Glu Ser Pro Thr Arg Arg Tyr Asp Asp Gly Trp Arg Arg Ser
    850                 855                 860

Ser Tyr Gln Asn Ile Tyr Gly Gln Gly Ala Ala Gln Phe Gln Ala Gln
865                 870                 875                 880

Pro Asp Leu Tyr Gly Lys Gln Val Ser Ala Thr Glu Trp Arg Gln Arg
                885                 890                 895

Leu Asp Thr His Glu Gln Gln Gln Arg Gln Arg Gln Arg Arg Ser Pro
            900                 905                 910

Phe Glu Asn Tyr Arg Gln Asp Val Glu Ile Asp Met Gln Lys Ala Arg
        915                 920                 925

Arg Asn Ser His Gly Asp Val Ile Asp Leu His Gly Thr Pro Asn Ser
    930                 935                 940

Ser Val Glu Glu Arg Phe Arg Arg Arg Gly Glu Pro Phe Ser Ala Glu
945                 950                 955                 960

Ser Gly Asp Asp Ser Ser Tyr Arg His Gln Gln Ile Met Ala Met Pro
                965                 970                 975

Ala Ala Gly Ser Ala Pro Ser Ala Lys Arg Tyr His Arg Arg Arg Ser
            980                 985                 990

Ser Glu Asp Ser Thr Ser Arg His  Gln Arg Trp Pro Ala  Asn Ile Glu
        995                 1000                1005

Glu Arg  Arg Ala Arg Arg Ala  Tyr Ser Pro Ala His  Asn Arg Pro
    1010                1015                1020
```

Glu Thr Ala Leu Thr Ser Tyr Ala Tyr Arg Ser Ser Ser His His
1025 1030 1035

Asn Leu Tyr Gln Pro Asn Gly Lys Ser Ser Lys Tyr Pro Pro Thr
1040 1045 1050

Tyr Gln Tyr Gly Asp Tyr Tyr His
1055 1060

<210> 10

<211> 332

<212> DNA

<213> Homo sapiens

<400> 10
agaattattg cacagggccg ggcgcggtgg ctcacgcctg taatcccagc actttgggag 60

gccgaggcgg gtggatcatg aggtcaggag atcgagacca tcctggctaa caaggtgaaa 120

ccccgtctct actaaaaata caaaaaatta gccgggcgcg gtggcgggcg cctgtagtcc 180

cagctactcg ggaggctgag gcaggagaat ggcgtgaacc cgggaagcgg agcttgcagt 240

gagccgagat tgcgccactg cagtccgcag tccgacctgg gcgacagagc gagactccgt 300

ctcaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 332

<210> 11

<211> 873

<212> DNA

<213> Homo sapiens

<400> 11
ccatgggaaa caaacagaga agaactctat gatctcttgg aaaccctgag gagactttct 60

gtcacctcca aggtgaagtt catcgtcttc aatccgtcct tgtatacatg gatcgatatg 120

cctcctctct gggagccccc ctgcacaact cctgcatcag tgcttgttcc tgctcttctt 180

ctcggcattc ctggtggcag attcactgat taacgtctgg atcactctca cagttgtgtc 240

cgtggagttg gagtgatagg tttcatgaca ttatggaaag tagaactgga ctgcatttct 300

gtgctatgct taatttatgg aattaattac acaattgaca attgtgctcc aatgttatcc 360

acatttgttc tgggcaagga tttcacaaga actaaatggg taaaaaatgc cctggaagtg 420

catggggtag ctattttaca gagttacctc tgctatatgt tggtctgatt cctcttgcag 480

ctgtgccttc aaatctgacc tgtacactgt tcaggtgctt gtttttaata gcattgtcac 540

cttctttcac tgcttgccat tttacctgtg atactgactt tcctgccacc ctctaagaaa 600

aaaaggaaag agaagaaaaa tcctgagaac cgggaggaaa ttgagtgtgt agaaatggta 660

gatatcgata gtacccgtgt ggttgaccaa attacaacag gtgatatgtc tgcttggata 720

tttcaccttа ggtttatcag aacaagagat ttgttatgaa acattaattc caaggtcttc 780

cctttaaaga ttgaacggtt ggcaaaaaaa aaaaaatggg gtttttgggg ggaaaggtta 840

agggggaagg tcgccttgaa gggaagggcc tag 873

<210> 12

<211> 346

<212> DNA

<213> Homo sapiens

<400> 12
gactgatcga cttcagacat catggaaact cgcactggat tccatttctt tgctatgctt 60

tatttatggg attcactaca caattgacca ctgtgctcca ctttcatcca catttgttct 120

aggcaaggat tttacaagaa ctaaatgggt ttaaaatccc ctggaagtgc atggggtagc 180

tattttacag agttacctct gttatattgt tggtctattt cctcttgcag ctgtcccttc 240

aaatctgacc tgtacactgt tcaggtgctt gtttttaata gcatttgtca ccttctttca 300

ctgctttgcc attttacctg tgatacttac tttcctgcca ccctcc 346

<210> 13

<211> 1692

<212> DNA

<213> Homo sapiens

<400> 13
atgctcaatg aagaaaaagt ggatttttcc attgctggtc atggattata tgggactttt 60

gaaatgttat cctcctggag gaaaactaga gaagaccaac atgttaaaga gagaactgca 120

gcagtctatg cagactccat gctctccttt tctctcacca ctgccatgta cctggtcacc 180

tttggcatag gggccagccc tttcacgaac attgaggcag ccaggatttt ctgctgcaat 240

tcctgtattg caatcttctt caactacctc tatgtactct cgttttatgg ttccagccta 300
gtgttcactg gctacataga aaacaattac cagcatagta tcttctgtag aaaagtccca 360
aagcctgagg cattgcagga gaagccggca tggtacaggt ttctcctgac ggccagattc 420
agtgaggaca cagctgaagg cgaggaagcg aacacttacg agagtcacct attggtatgt 480
ttcctcaaac gctattactg tgactggata accaacacct atgtcaagcc ttttgtagtt 540
ctcttttacc ttatttatat ttcctttgcc ttaatgggct atctgcaggt cagtgaaggg 600
tcagacctta gtaacattgt agcaaccgcg acacaaacca ttgagtacac tactgcccag 660
caaaagtact tcagcaacta cagtcctgtg attgggtttt acatatatga gtctatagaa 720
tactggaaca ctagtgtcca agaagatgtt ctagaataca ccaaggggtt tgtgcggata 780
tcctggtttg agagctattt aaattacctt cggaaactca atgtatccac tggcttgcct 840
aagaaaaatt tcacagacat gttgaggaat tcctttctga aagcccctca attttcacat 900
tttcaagagg acatcatctt ctctaaaaaa tacaatgatg aggtcgatgt agtggcctcc 960
agaatgtttt tggtggccaa gaccatggaa acaaacagag aagaactcta tgatctcttg 1020
gaaaccctga ggagactttc tgtcacctcc aaggtgaagt tcatcgtctt caatccgtcc 1080
tttgtataca tggatcgata tgcctcctct ctgggagccc ccctgcacaa ctcctgcatc 1140
agtgctttgt tcctgctctt cttctcggca ttcctggtgg cagattcact gattaacgtc 1200
tggatcactc tcacagttgt gtccgtggag tttggagtga taggtttcat gacattatgg 1260
aaagtagaac tggactgcat ttctgtgcta tgcttaattt atggaattaa ttacacaatt 1320
gacaattgtg ctccaatgtt atccacattt gttctgggca aggatttcac aagaactaaa 1380
tgggtaaaaa atgccctgga agtgcatggg gtagctattt tacagagtta cctctgctat 1440
attgttggtc tgattcctct tgcagctgtg ccttcaaatc tgacctgtac actgttcagg 1500
tgcttgtttt taatagcatt tgtcaccttc tttcactgct ttgccatttt acctgtgata 1560
ctgactttcc tgccaccctc taagaaaaaa aggaaagaga agaaaaatcc tgagaaccgg 1620
gaggaaattg agtgtgtaga aatggtagat atcgatagta cccgtgtggt tgaccaaatt 1680
acaacagtgt ga 1692

<210> 14

<211> 563

<212> PRT

<213> Homo sapiens

<400> 14

```
Met Leu Asn Glu Glu Lys Val Asp Phe Ser Ile Ala Gly His Gly Leu
1               5                  10                  15

Tyr Gly Thr Phe Glu Met Leu Ser Ser Trp Arg Lys Thr Arg Glu Asp
            20                  25                  30

Gln His Val Lys Glu Arg Thr Ala Ala Val Tyr Ala Asp Ser Met Leu
        35                  40                  45

Ser Phe Ser Leu Thr Thr Ala Met Tyr Leu Val Thr Phe Gly Ile Gly
    50                  55                  60

Ala Ser Pro Phe Thr Asn Ile Glu Ala Ala Arg Ile Phe Cys Cys Asn
65                  70                  75                  80

Ser Cys Ile Ala Ile Phe Phe Asn Tyr Leu Tyr Val Leu Ser Phe Tyr
                85                  90                  95

Gly Ser Ser Leu Val Phe Thr Gly Tyr Ile Glu Asn Asn Tyr Gln His
            100                 105                 110

Ser Ile Phe Cys Arg Lys Val Pro Lys Pro Glu Ala Leu Gln Glu Lys
        115                 120                 125

Pro Ala Trp Tyr Arg Phe Leu Leu Thr Ala Arg Phe Ser Glu Asp Thr
    130                 135                 140

Ala Glu Gly Glu Glu Ala Asn Thr Tyr Glu Ser His Leu Leu Val Cys
145                 150                 155                 160

Phe Leu Lys Arg Tyr Tyr Cys Asp Trp Ile Thr Asn Thr Tyr Val Lys
                165                 170                 175

Pro Phe Val Val Leu Phe Tyr Leu Ile Tyr Ile Ser Phe Ala Leu Met
            180                 185                 190

Gly Tyr Leu Gln Val Ser Glu Gly Ser Asp Leu Ser Asn Ile Val Ala
        195                 200                 205

Thr Ala Thr Gln Thr Ile Glu Tyr Thr Thr Ala Gln Gln Lys Tyr Phe
    210                 215                 220
```

```
Ser Asn Tyr Ser Pro Val Ile Gly Phe Tyr Ile Tyr Glu Ser Ile Glu
225                 230                 235                 240

Tyr Trp Asn Thr Ser Val Gln Glu Asp Val Leu Glu Tyr Thr Lys Gly
                245                 250                 255

Phe Val Arg Ile Ser Trp Phe Glu Ser Tyr Leu Asn Tyr Leu Arg Lys
            260                 265                 270

Leu Asn Val Ser Thr Gly Leu Pro Lys Lys Asn Phe Thr Asp Met Leu
        275                 280                 285

Arg Asn Ser Phe Leu Lys Ala Pro Gln Phe Ser His Phe Gln Glu Asp
    290                 295                 300

Ile Ile Phe Ser Lys Lys Tyr Asn Asp Glu Val Asp Val Val Ala Ser
305                 310                 315                 320

Arg Met Phe Leu Val Ala Lys Thr Met Glu Thr Asn Arg Glu Glu Leu
                325                 330                 335

Tyr Asp Leu Leu Glu Thr Leu Arg Arg Leu Ser Val Thr Ser Lys Val
            340                 345                 350

Lys Phe Ile Val Phe Asn Pro Ser Phe Val Tyr Met Asp Arg Tyr Ala
        355                 360                 365

Ser Ser Leu Gly Ala Pro Leu His Asn Ser Cys Ile Ser Ala Leu Phe
    370                 375                 380

Leu Leu Phe Phe Ser Ala Phe Leu Val Ala Asp Ser Leu Ile Asn Val
385                 390                 395                 400

Trp Ile Thr Leu Thr Val Val Ser Val Glu Phe Gly Val Ile Gly Phe
                405                 410                 415

Met Thr Leu Trp Lys Val Glu Leu Asp Cys Ile Ser Val Leu Cys Leu
            420                 425                 430

Ile Tyr Gly Ile Asn Tyr Thr Ile Asp Asn Cys Ala Pro Met Leu Ser
        435                 440                 445

Thr Phe Val Leu Gly Lys Asp Phe Thr Arg Thr Lys Trp Val Lys Asn
    450                 455                 460
```

Ala Leu Glu Val His Gly Val Ala Ile Leu Gln Ser Tyr Leu Cys Tyr
465 470 475 480

Ile Val Gly Leu Ile Pro Leu Ala Ala Val Pro Ser Asn Leu Thr Cys
485 490 495

Thr Leu Phe Arg Cys Leu Phe Leu Ile Ala Phe Val Thr Phe Phe His
500 505 510

Cys Phe Ala Ile Leu Pro Val Ile Leu Thr Phe Leu Pro Pro Ser Lys
515 520 525

Lys Lys Arg Lys Glu Lys Lys Asn Pro Glu Asn Arg Glu Glu Ile Glu
530 535 540

Cys Val Glu Met Val Asp Ile Asp Ser Thr Arg Val Val Asp Gln Ile
545 550 555 560

Thr Thr Val

<210> 15

<211> 296

<212> DNA

<213> Homo sapiens

<400> 15
ctcctattct tgattaatgg aatcgactac acaactgaca actctgctcc actgttatcc 60

acatttcttc taggcaagga tttcaccaga actaaatggg ttaaaaatcc cctggaagtg 120

catggggtag ctattttaca gagttacctc tgttatattg ttggtctatt tcctcttgca 180

gctgtccctt caaatctgac ctgtacactg ttcaggtgct tgtttttaat agcatttgtc 240

accttctttc actgctttgc cattttacct gtgatactta ctttcctccc accctc 296

<210> 16

<211> 290

<212> DNA

<213> Homo sapiens

<400> 16
gcttaattta tggaattaac tacacaattg acaactctgc tccattgtct tccaaatttg 60

ttccaggcaa ggattttaca agaactaaat gggttaaaaa tgccctggaa gtgcatgggg 120

tagctatttt acagagttac ctctgttata ttgttggtct atttcttctt gcagctgtgc 180

cttcaaatct gacctgtaca ctgttcaggt gcttgttttt aatagcattt gtcaccttct 240

ttcactgctt tgccatttta cctgtgatac ttactttcct gccaccctcg 290